

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOHAMED EL-BACHIR EL-IBRAHIMI - BORDJ BOU ARRERIDJ



THESE DE DOCTORAT

Présentée à la Faculté des Sciences et de la Technologie
Département d'Electronique

Pour l'obtention du Diplôme de
DOCTORAT EN SCIENCE

Spécialité

ELECTRONIQUE

THEME

***Extraction et sélection des paramètres pertinents
appliquées à l'identification des appareils électriques***

Par : Mr. GHEZZALI Fatah

Soutenue le : **09/06/2022** devant le jury composé de

Dr. FLISSI Moustapha	Président	MCA	Université de BBA
Dr. HACINE-GHARBI Abdenour	Encadreur	MCA	Université de BBA
Pr. ROUABAH Khaled	Co-Encadreur	Professeur	Université de BBA
Dr. LADJAL Mohamed	Examineur	Professeur	Université de M'sila
Dr. MESLOUB Ammar	Examineur	MCA	EMP, Bordj El Bhri, Alger
Dr. OUALI Mohammed Assam	Examineur	MCA	Université de M'sila

Année Universitaire 2021/2022

Remerciements

Avant tout, je remercie Dieu de m'avoir donné la force pour réaliser ce présent travail.

*Je tiens à exprimer d'abord ma très profonde gratitude à Monsieur **Abdenour Hacine-Gharbi**, MCA à l'université Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj, qui n'a ménagé aucun effort pour me prendre en charge pour la réalisation de ce travail. Sa clairvoyance, sa générosité, sa gentillesse, ses connaissances, le temps qu'il m'a dispensé, et sa grande disponibilité dont il a fait preuve, m'ont énormément facilité ma tâche. Les mots risquent de me manquer pour lui exprimer toute ma gratitude et ma reconnaissance.*

*Je tiens aussi à adresser mes remerciements les plus profonds à monsieur **Philippe Ravier**, maître de conférences à l'université d'Orléans - France- pour ses aides précieuses qui ont judicieusement éclairé mon chemin vers l'aboutissement et la concrétisation de ma thèse.*

*J'exprime également mon sincère respect et mon appréciation à l'âme du professeur, **Mohamadi Tayeb**, qui était l'encadreur de ce sujet jusqu'au jour où nous sommes séparés en été 2020, et à cet égard, je ne peux qu'ajouter à son âme pure les plus hautes expressions de respect et d'appréciation. *إنا لله و إنا إليه راجعون**

*Notre reconnaissance va aussi à Monsieur **Rouabah Khaled** Professeur à l'université Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj. Veuillez bien trouver l'expression de ma vive gratitude, ses conseils avisés, sa prise en charge et sa générosité m'ont été de grande utilité pour mon travail de recherche.*

*J'adresse mes très sincères remerciements à Monsieur **Flissi Moustapha** pour l'intérêt qu'il a accordé à ce travail en acceptant de le juger et de présider le jury.*

Je tiens à remercier les membres du jury qui m'ont fait l'extrême honneur de participer à l'examen de ce travail.

*A toute ma famille,
et aux gens que j'aime !*

Résumé

De nos jours, la demande croissante de l'énergie électrique domestique oblige les fournisseurs d'électricité d'améliorer la gestion de distributions énergétiques par l'installation des compteurs intelligents. Ces derniers incorporent dans leur traitement une tâche d'identification des appareils électriques en temps réel en identifiant les classes d'appareils en état de fonctionnement à partir du signal de courant électrique mesuré au niveau d'un compteur électrique placé à l'intérieur d'un boîtier domestique standard installé à l'entrée de la maison.

Notre objectif, dans le cadre de cette thèse de Doctorat, est de proposer un nouveau système d'identification des appareils électriques domestiques. Notre première contribution consiste à proposer un système d'identification basé sur l'utilisation des paramètres statistiques des harmoniques et l'application du classificateur KNN combiné avec la méthode de règle de vote. Les résultats obtenus ont montré que l'extraction de 500 paramètres, basée sur l'estimation de la moyenne statistique et de l'écart type, combinée avec la classification KNN et la stratégie de la règle de vote, donne le meilleur Taux de classification CR qui est de 94.97%. Les résultats ont montré également de bonnes performances en utilisant d'autres métriques telles que la sensibilité, la précision et le F-score. Une procédure de sélection de paramètres, basée sur la stratégie JMI, montre que l'utilisation d'un sous-ensemble de cinq paramètres est suffisante pour expliquer les différentes classes d'appareils.

Notre deuxième contribution, consiste à réduire la dimensionnalité en utilisant une représentation compacte des paramètres (appelée DWE) qui est basée sur l'estimation de la moyenne et de l'écart type de l'énergie calculée à chaque niveau de décomposition dyadique de l'analyse par ondelettes. Deux descripteurs appelés LWE et WCC sont extraits également de cette analyse en appliquant respectivement le logarithme de l'Énergie totale puis la transformée en cosinus discrète. Les résultats obtenus montrent que le descripteur WCC donne un CR maximal de 98.13%. De plus, la combinaison et la sélection des paramètres du descripteur WCC avec les paramètres statistiques du logarithme de l'énergie totale (descripteur LOG_E) améliore le taux de classification qui atteint jusqu'à 98.51 en utilisant seulement 5 paramètres.

Mots clés : Identification des appareils électriques, analyse harmoniques, extraction de paramètres statistiques, k plus proches voisins, méthode de règle de vote, approche de sélection de paramètres de type "*Filters*", information mutuelle, analyse par ondelettes discrètes; approche de sélection de paramètres de type "*wrappers*".

Abstract

Nowadays, the increasing energy consumption compels the electricity suppliers to implement a new strategy of electric power distribution. From this perspective, the installation of smart meters is essential. The latter ones integrate, into their processing process, an important real-time function which is the electrical appliance identification from measurements of current signal at the electrical meter that is placed inside a standard household electric meter box at the entrance of the house.

The main objective of this doctoral work is to propose a new electrical appliances identification system. Our first contribution consists in proposing an identification system based on the use of statistical harmonics features and the application of the k-NN classifier combined with the voting rule method. The obtained results show that the extraction of 500 features based on estimation of statistical mean and standard deviation, combined with KNN classification and voting rule strategy, gives the best classification rate CR which is equal to 94.97%. The results have also shown good performance using other metrics such as sensitivity, accuracy and F-score. A features selection procedure, based on the JMI strategy, have shown that only a subset of five features is sufficient to explain the different electrical appliance classes.

Our second contribution consists in reducing the dimensionality by using a compact representation of features. The last representation, called DWE, is based on the estimate of both mean and standard deviation of the energy calculated at each level of dyadic decomposition of the wavelet analysis. Two descriptors, called LWE and WCC, are also extracted from this analysis by applying respectively the logarithm of the total energy and the discrete cosine transform. The obtained results have shown that the WCC descriptor gives a maximum CR of 98.13%. Moreover, the combination and selection of the WCC descriptor features with the statistical features of the logarithm of the total energy (LOG_E descriptor) improves the classification rate up to 98.51% using only 5 features.

Key words: Electrical appliance identification, harmonics analysis, statistical features extraction, k-Nearest Neighbors, voting rules method, Filters features selection approach, mutual information, discrete wavelet analysis; wrapper feature selection approach.

ملخص الأطروحة

في الوقت الحاضر، تتطلب الزيادة في استهلاك الطاقة من موردي الكهرباء تنفيذ إستراتيجية جديدة لتوزيع الطاقة الكهربائية. من هذا المنظور، فإن تركيب العدادات الذكية أمر ضروري. تشمل هذه العدادات على وظيفة مهمة في عملية المعالجة الخاصة بها، وهي تحديد فئة الجهاز في الوقت الفعلي من القياسات التي تم إجراؤها على مستوى العداد الكهربائي الرئيسي.

هدفنا هو اقتراح نظام جديد لتحديد الأجهزة الكهربائية. تتمثل مساهمتنا الأولى في اقتراح نظام تحديد يعتمد على استخدام ميزات التوافقيات الإحصائية وتطبيق مصنف KNN مع طريقة قاعدة التصويت. تظهر النتائج أن استخراج 500 ميزة بناءً على تقدير المتوسط الإحصائي والانحراف المعياري للميزات مركبة مع المصنف KNN و طريقة قاعدة التصويت تعطي أفضل نسبة للتصنيف تقدر بـ 94.97%. كما أظهرت النتائج أداء جيداً باستخدام مقاييس أخرى مثل الحساسية و الدقة و بما يعرف بـ F-score. بينت عملية اختيار الميزات باستعمال إستراتيجية JMI أن استخدام مجموعة فرعية مكونة من خمسة ميزات كافية لشرح مختلف فئات الأجهزة.

تتمثل مساهمتنا الثانية في تقليل الأبعاد باستخدام تمثيل مضغوط للميزات تسمى DWE والتي تستند إلى تقدير المتوسط الإحصائي والانحراف المعياري للطاقة المحسوبة في كل مستوى من مستويات التحليل الثنائي لتحليل الموجات. يتم أيضاً استخلاص اثنين من الواصفات تسمى LWE و WCC من هذا التحليل من خلال تطبيق اللوغاريتم ثم تحويل جيب التمام المتقطع على التوالي. وتبين النتائج أن واصف WCC يعطي نسبة تصنيف قصوى تبلغ 98.13%. علاوة على ذلك، فإن دمج واختيار ميزات واصف WCC مع الميزات الإحصائية للوغاريتم للطاقة الإجمالية (واصف LOG_E) يحسن معدل التصنيف حتى 98.51% باستخدام 5 ميزات فقط.

الكلمات المفتاح : تحديد الأجهزة الكهربائية ، تحليل التوافقيات ، استخراج الخصائص الإحصائية ، المصنف KNN ، طريقة قواعد التصويت ، مقارنة المرشحات لاختيار الميزات ، المعلومات المتبادلة ، تحليل الموجات المتقطع ، مقارنة wrappers لاختيار الميزات

Table des matières

Remerciements	
résumé	
liste de Figures	
liste de tableaux	
INTRODUCTION GENERALE	1
CHAPITRE I: GENERALITES SUR L'IDENTIFICATION DES APPARELS ELECTRIQUES	
I-1 Introduction.....	5
I-2 Objectifs d'identification des appareils électriques.....	6
I-3 Identification des appareils électriques dans l'approche NILM	7
I-4 Principe de fonctionnement d'un système IAED	9
I-5 Etat de l'art sur les systèmes d'identification des appareils électriques	11
I-6 Algorithmes de classification	13
I-6-1- Classificateur Machines à Vecteurs de Support SVM	14
I-6-2 Classificateur K-plus proche voisin (KNN: K-Nearest Neighbours Classifier)	16
I-6-2-1 Quelques exemples de distances	17
I-6-2-1-a Distance euclidienne	17
I-6-2-1-b Distance de Cityblock	17
I-6-2-1-c Distance Cosinus	17
I-6-2-1-d Distance Corrélacion	18
I-6-2-1-e Distance de Hamming	18
I-6-2-2 L'algorithme de base du classificateur KNN	18
I-6-3 Modèles de Markov cachés (HMM: Hidden Markov Models)	19
I-7. Mesures des performances d'un système d'identification	20
I-7.1. Taux de classification TC (RC : classification ratio)	20
I-7-2. Matrice de confusion	21
I-7-3 Mesures statistiques	22

I-8 Approche de classification proposée pour la conception d'un système IAED	22
I-8-1 Stratégie de règle de vote	23
I-9 conclusion	23

CHAPITRE II: METHODES D'EXTRACTION DE PARAMETRES

II-1 Introduction	24
II-2 Etat de l'art sur l'extractions des paramètres	24
II-2-1 Puissance réelle	25
II-2-2 Paramètres PQ	25
II-2-3 Paramètres PQ combinés avec les paramètres transitoires macroscopiques ..	25
II-2-4 Paramètres PQ combinés avec le courant I et la tension V	26
II-2-5 Paramètres PQ combinés avec les paramètres transitoires macroscopiques et les paramètres harmoniques	26
II-2-6 Paramètres P-Q combinés avec les harmoniques	26
II-2-7 Enveloppe spectrale	26
II-2-8 Paramètres basés sur la transformée en ondelettes	27
II-2-9 Trajectoires V-I (tension - courant)	28
II-2-10 Courant inactif	28
II-2-11 Paramètres non conventionnelles	29
II-3 Analyse harmonique	29
II-3-1. Décomposition en Série de Fourier (DSF)	30
II-3-2 Application de l'analyse harmonique dans les systèmes IAED	31
II-3-3 Approche d'extraction proposée	32
II-4. Transformée de Fourier à court terme STFT	34
II-5 Analyse par Ondelette	35
II-5-1 Historique des ondelettes	35
II-5-2 Transformée en ondelette continue (TOC)	36
II-5-3 Transformée en ondelette discrète	40

II-5-4 Application de l'analyse en ondelette dans les systèmes IAED	42
II-6 Conclusion	44

**CHAPITRE III: SELECTION DES PARAMETRES PERTINENTS POUR LA
CLASSIFICATION**

III-1 Introduction	45
III-2 Définitions de la sélection des paramètres	46
III-3 Classification des méthodes de Sélection	48
III-3-1 Méthodes de sélection " <i>Filters</i> "	49
III-3-2 Méthodes " <i>Wrappers</i> "	51
III-3-3 Méthodes de sélection hybrides	52
III-4 Méthodes de sélection fondée sur l'information mutuelle (IM)	53
III-4-1- Entropie et Information mutuelle	55
III-4-1-1- Entropie (information moyenne)	55
III-4-1-2- Information mutuelle	56
III-4-1-3 Information mutuelle d'ordre supérieur	57
III-4-2 Estimation de l'entropie et de l'information mutuelle	59
III-4-2-1 Formules du choix optimal du nombre de <i>bins</i>	59
III-4-3 Stratégies de sélection " <i>Filters</i> " basée sur IM	60
III-5 Application de la sélection des paramètres pertinents dan les systèmes IAED	62
III-6 Conclusion	63

CHAPITRE IV: SYSTEME PROPOSE : IMPLIMENTATION ET RESULTATS

IV-1 Introduction	65
IV-2 Architecture globale des systèmes proposés	66
IV-2-1 Base de donnée PLAID	67
IV-2-2 Préparation de données	68
IV-2-3 Extraction des paramètres statistiques	69
IV-2-3-1 Prétraitement	69

IV-2-3-2 Fenêtrage	69
IV- 2-3-3 Analyse du signal à court terme	71
IV- 2-3-3-1 Analyse harmonique à court terme	71
IV- 2-3-3-2 Analyse par ondelettes à court terme	71
IV-2-3-4 Estimation des paramètres statistiques	72
IV-2-4 Classification KNN	73
IV-2-5 Application de la stratégie de règle de vote	74
IV-2-6 Sélection des paramètres statistiques pertinents	74
IV-3 Expériences et résultats	75
IV-3-1 Système IAED basé sur l'analyse harmonique	75
IV-3-1-1 Combinaison optimale des paramètres statistiques	75
IV-3-1-2 Configuration optimale de classificateur KNN	77
IV-3-1-3 Nombre optimal de vecteurs statistiques de vote	78
IV-3-1-4 Effet de la taille des bases de données d'apprentissage et de test sur les performances d'identification	79
IV-3-1-5 Evaluation approfondie des performances	80
IV-3-1-6 Sélection des paramètres statistiques STFS	82
IV-3-2 Système IAED basé sur l'analyse par ondelettes	84
IV-3-2-1 Comparaison des performances du système IAED	85
IV-3-2-2 Durée optimale de la fenêtre d'analyse	87
IV-3-2-3 Choix de l'ondelette mère et du niveau de décomposition dyadique	88
IV-3-2-4 Évaluation approfondie des performances	89
IV-3-4 Résultats de la sélection des paramètres	90
IV-4 Conclusion	93
Conclusion Générale et perspectives.....	95

Bibliographie

Liste de Figures

CHAPITRE I: GENERALITES SUR L'IDENTIFICATION DES APPARELS ELECTRIQUES

Figure I-1 Composants principaux d'un système de surveillance de la charge des appareils électriques	5
Figure I-2 Classification des méthodes de surveillance de la consommation d'énergie électrique	8
Figure I-3 Schéma synoptique général du processus d'identification IAED.	10
Figure I-4: Schéma synoptique de la conception d'un système IAED	10
Figure I-5 Hyper-plan à marge maximale pour un SVM formé avec des échantillons de deux classes.	16
Figure I-6 : Exemple d'un Modèle de Markov gauche-droite	19

CHAPITRE II: METHODES D'EXTRACTION DE PARAMETRES

Figure II-1 Schéma du système d'identification proposé dans [HYPERLINK \l "Nai161" 1]	31
Figure II-2 Plan multi-résolution temps-fréquence	36
Figure II-3 Ondelette de Morlet avec différentes échelles ($s=0.5, s=1, s=2, s=4$).....	38
Figure II-4: Ondelette de Morlet avec différents facteurs de translation ($\tau = -0.5, \tau = 0$ et $\tau = 0.5$)	39
Figure II-5 Quelques familles d'ondelettes	39
Figure II-6 Schéma de la décomposition en ondelettes	41
Figure II-7 Processus d'extraction des paramètres DWE/LWE/WCC avec fenêtrage de Hamming.....	42

CHAPITRE III: SELECTION DES PARAMETRES PERTINENTS POUR LA CLASSIFICATION

Figure III-1 Processus général de sélection des paramètres.....	47
Figure III-2 Algorithme de sélection <i>Wrapper</i>	52
Figure III-3 Diagramme reliant l'IM et l'entropie	57
Figure III-3 Diagramme reliant l'IM et l'entropie	58

CHAPITRE IV: SYSTEME PROPOSE : IMPLIMENTATION ET RESULTATS

Figure IV-1 Architecture globale des Systèmes IAED proposés	67
Figure IV-2: Etape d'extraction des paramètres statistiques	69
Figure IV-3 Etape de fenêtrage	70
Figure IV-4: Analyse harmonique à court terme de la fenêtre d'analyse	71
Figure IV-5: Analyse par ondelettes de la fenêtre d'analyse	72
Figure IV-6: Etape d'extraction des vecteurs de paramètres statistiques	72
Figure IV-7: Etape de classification KNN	73
Figure IV-8: Application de la stratégie de la règle de vote	74
Figure IV-9: CR (%) pour la sélection de paramètres par la stratégie JMI avec les trois répartitions de la base de données.	83

liste de tableaux

Tableau I-1: Etat de l'art récapitulatif sur les techniques de classification utilisées dans systèmes IAED.	13
Tableau I-2 : Description de la matrice de confusion dans la classification binaire	21
Tableau III-1: Différentes méthodes de sélection " <i>Filters</i> "	51
Tableau IV-1: Récapitulatif sur la base de données PLAID	67
Tableau IV-2: Nombre d'instances pour une répartition 50%/50% de la base de données PLAID.	68
Tableau IV-3 Meilleure combinaison de paramètres statistiques	76
Tableau IV-4: Configuration optimale de classificateur KNN	78
Tableau IV-5: Nombre optimal de vecteurs statistiques (par seconde) donné par la méthode des règles de vote, en utilisant 250 harmoniques.	79
Tableau IV-6: Effet de la répartition de la base de données sur les résultats de CR	80
Tableau IV-7: Sensibilité S , précision P et F-score Fs de chaque type d'appareil électrique pour différentes répartitions de la base de données	80
Tableau IV-8: Table de la matrice de confusion	81
Tableau IV-9: Taux de classification CR (%) des 10 premiers paramètres sélectionnés comparés aux 500 paramètres	84
Tableau IV-10: Comparaison des performances en terme de CR en % des différents types de descripteurs pour différentes valeurs du nombre de vecteurs statistiques.	86
Tableau IV-11 : Taux de classification CR pour différents descripteurs et pour différentes durées de la fenêtre d'analyse.....	87
Tableau IV-12: CR% maximal obtenu pour chaque ordre N de la famille Daubechies et son niveau de décomposition correspondant	88
Tableau IV-13 : CR% maximal obtenu pour chaque ordre N de la famille Coiflet et son niveau de décomposition correspondant	88

Tableau IV-14 : CR% maximal obtenu pour chaque ordre N de la famille Symlet et son niveau de décomposition correspondant	88
Tableau IV-15: Sensibilité S, précision P et F-score Fs de chaque type d'appareil électrique	89
Tableau IV-16 Matrice de confusion obtenue avec la configuration optimale	90
Tableau IV-17: CR% en fonction des paramètres sélectionnés, Sel est le numéro du paramètre sélectionné et Feat est le nom du paramètre sélectionné à l'itération j.	91
Tableau IV-18: CR% en fonction des paramètres sélectionnés avec concaténation de chaque descripteur (DWE, LWE ou WCC) avec le descripteur LOG_E.	92

Acronymes

<i>IAED</i>	<i>Identification des Appareils Electriques Domestiques</i>
<i>NILM</i>	<i>Non-Intrusive Load Monitoring</i>
<i>IM</i>	<i>Information Mutuelle</i>
<i>P</i>	<i>Précision</i>
<i>S</i>	<i>Sensibilité</i>
<i>STFT</i>	<i>Short-term Fourier Transform</i>
<i>FT:</i>	<i>Fourier Transform</i>
<i>FFT:</i>	<i>Fast Fourier Transforms</i>
<i>CWT</i>	<i>Continues Wavelet Transform</i>
<i>DWT</i>	<i>Discrete wavelet Transform</i>
<i>pdf</i>	<i>Probability Density Function</i>
<i>TFTD</i>	<i>Transformée de Fourier à temps discret</i>
<i>CR</i>	<i>Classification rate</i>
<i>DCT</i>	<i>Discrete cosine transform</i>
<i>DWE</i>	<i>Discrete wavelet energy</i>
<i>HMM</i>	<i>Hidden Markov Models</i>
<i>KNN</i>	<i>K nearest neighbour</i>
<i>LOG_E</i>	<i>Logarithm energy</i>
<i>LWE</i>	<i>Log wavelet energy</i>
<i>STFS</i>	<i>Short-term Fourier series</i>
<i>SVM</i>	<i>Support vector machine</i>
<i>WCC</i>	<i>Wavelet cepstral coefficients</i>
<i>DSF</i>	<i>Décomposition en Série de Fourier</i>
<i>TO</i>	<i>Transformé en ondelette</i>
<i>SRV</i>	<i>Stratégie de Règle de Vote</i>
<i>DFT</i>	<i>Discret Fourier Transform</i>
<i>TOC</i>	<i>Transformé en ondelette continue</i>
<i>TOD</i>	<i>Transformée en ondelette discrète</i>

INTRODUCTION
GENERALE

INTRODUCTION GENERALE

Durant ces dernières années, le développement extraordinaire de nouvelles technologies en matière de croissance économique, du respect de l'environnement et également de la consommation énergétique domestique ne cesse de s'accroître rapidement. Dans cette optique, la mise en place d'un système de contrôle de la consommation de l'énergie a devenu une nécessité, dans lequel plusieurs travaux de recherche ont proposé l'installation des compteurs intelligents («smart meter» en anglais) dans les réseaux électriques domestiques, afin que les clients puissent optimiser leurs factures énergétiques, et mieux maîtriser leurs consommations électriques. De plus, l'utilisation des compteurs intelligents va permettre aux fournisseurs d'énergie de mieux gérer leurs distributions énergétiques (pour des opérations de délestage ou de facturation).

L'évolution profonde des enjeux énergétiques a accéléré l'émergence de la thématique des réseaux intelligents «smart grids». Sous ce terme générique se dessinerait la prochaine génération de réseaux électriques visant à répondre mieux aux problématiques associées à la production et à la distribution d'électricité. Cette situation nous a conduit à envisager les différentes manières de production et d'utilisation de l'électricité. L'optimisation de la demande d'énergie est un processus de gestion de la demande d'énergie en fonction des ressources de production disponibles afin de maintenir un équilibre entre l'offre et la demande. L'identification des appareils électriques domestiques (par exemple les systèmes d'éclairage et de chauffage, les machines à laver, systèmes de production d'eau chaude sanitaire et autres), est également intégrée dans les compteurs intelligents pour répondre aux questions : qui est en train de consommer ? quand ? et combien consomme-t-il ? Cette identification peut s'effectuer de manière intrusive (à l'intérieur de l'habitat) ou non-intrusive (à l'extérieure de l'habitat). Ce présent travail s'intéresse à la méthode non-intrusive NILM (Non-Intrusive Load Monitoring).

Le cadre général de la méthode NILM commence à partir des mesures faites sur l'entrée de la consommation totale d'électricité pour finalement la désagréger en contributions individuelles de chaque appareil. Un système NILM peut être considéré comme un système de reconnaissance de formes qui exige une étape d'extraction des

paramètres caractérisant les différentes classes de formes et une étape de classification qui permet d'identifier la classe du signal d'entrée du système en se basant sur des modèles des classes entraînés durant la phase d'apprentissage. Plus particulièrement, un système d'Identification des Appareils Electriques Domestiques (IAED) consiste à identifier la classe d'appareil électrique à partir du signal électrique. La conception d'un tel système nécessite une phase d'apprentissage pour modéliser les différentes classes d'appareils et une phase de test pour évaluer les performances du système d'identification en se basant sur des algorithmes de classification. Ces deux phases exigent une étape d'extraction de paramètres permettant de convertir chaque signal électrique en une séquence de vecteurs de paramètres dont chaque vecteur représente une fenêtre d'analyse.

De nombreux modèles de classes ont été proposés dans plusieurs travaux de recherches, à titre d'exemple, on peut citer les modèles de Markov cachés (HMM) [1], les machines à vecteurs de supports (SVM) [2], les réseaux de neurones (ANN) [3], les k plus proches voisins (KNN) [4]. Autrement dit, un système IAED, basé sur la méthode HMM, consiste à modéliser chaque classe d'appareils durant la phase d'apprentissage et à classifier chaque séquence de vecteurs de paramètres représentant un signal d'entrée de test en une classe d'appareils [1]. Cependant, d'autres algorithmes de classification tels que le KNN, le SVM, l'ANN classifient chaque vecteur de paramètres et appliquent ensuite la règle de vote sur la séquence d'indices de classes pour classifier la séquence de vecteurs de paramètres qui représente le signal d'entrée de test. Pratiquement, la méthode de classification KNN est la plus utilisée dans le domaine de la reconnaissance de formes vue sa simplicité. Cependant, cette méthode souffre du problème de complexité en termes d'encombrement mémoire dans la phase d'apprentissage et du temps de calcul dans la phase de test.

Dans notre travail, on s'intéresse à minimiser cette complexité en réduisant le nombre de vecteurs de paramètres de chaque signal ainsi que la dimension de chaque vecteur. La réduction du nombre de vecteurs peut s'effectuer en extrayant un vecteur de paramètres statistiques à partir d'une sous-séquence de vecteurs de paramètres instantanés. La réduction de la dimension de chaque vecteur de paramètres peut être

réalisée en appliquant les méthodes de sélection de paramètres. En effet, cette réduction permet d'améliorer la précision du système d'identification en minimisant le phénomène de la malédiction et en réduisant le coût de calcul et l'encombrement mémoire.

Subséquentement, notre travail de thèse consiste à concevoir un système IAED basé sur le classificateur KNN combiné avec la stratégie de règle de vote. Ce travail a pour objectif d'améliorer les performances du système du point de vue complexité et précision en proposant une méthode d'extraction de paramètres statistiques ainsi que l'application des algorithmes de sélection de paramètres pertinents.

Notre première contribution consiste à proposer un système d'identification basé sur l'extraction des paramètres statistiques des coefficients harmoniques et l'application du classificateur KNN combiné avec la méthode de règle de vote [[HYPERLINK \l "Gha192" 5](#)]. L'étape d'extraction des paramètres découpe en premier lieu le signal du courant électrique en une séquence de fenêtres chevauchées. En deuxième lieu, elle transforme chaque fenêtre d'analyse en un vecteur de paramètres STFS calculés à partir du module des coefficients harmoniques, puis elle estime les paramètres statistiques (moyenne, écart type, asymétrie et l'aplatissement) sur chaque sous-séquence de N vecteurs de paramètres STFS. Ainsi, cette étape permet de convertir chaque signal en une séquence de vecteurs de paramètres statistiques. Cependant, l'application de cette méthode d'extraction sur des signaux échantillonnés avec une grande fréquence d'échantillonnage (30 KHz dans le cas de la base de données utilisée PLAID) conduit à l'extraction de vecteurs de grandes dimensions exigeant alors un grand espace mémoire et un temps de calcul considérable. Afin de remédier ce problème, on propose d'appliquer des algorithmes de sélection de paramètres pertinents de type "Filters" basés sur le critère de maximisation de l'information mutuelle.

Notre deuxième contribution, consiste à réduire la dimensionnalité en utilisant une représentation plus compacte des fenêtres d'analyse, basée sur le calcul de l'énergie à chaque niveau de décomposition en ondelettes discrètes (DWE: Discrete Wavelet Energy). Deux autres descripteurs appelés LWE et WCC sont également extraits en appliquant respectivement d'une façon consécutive le logarithme et la transformée DCT sur le descripteur DWE6}], [[HYPERLINK \l "Gha21" 7](#)]. La sélection de paramètres

pertinents pour ce type de descripteurs est effectuée en appliquant l'approche de sélection de type 'Wrappers', dont la mesure de pertinence est basée sur le taux de classification [7].

Le manuscrit de la thèse contient quatre chapitres :

- Le premier chapitre donne des généralités sur les systèmes d'identification des appareils électriques et leur fonctionnement. Un état de l'art est donné sur les classificateurs et les différentes métriques d'évaluation des performances des système IAED.
- Dans le deuxième chapitre, nous allons dresser un état de l'art sur les méthodes d'extraction de paramètres utilisées dans les systèmes IAED. Plus particulièrement, l'analyse harmonique et l'analyse par ondelettes et leurs applications dans les systèmes IAED seront décrites .
- Le troisième chapitre présente un état de l'art sur les méthodes de sélection des paramètres pertinents appliquées à l'identification des appareils électriques.
- Le quatrième chapitre est consacré aux systèmes proposés et leurs architectures. Différentes expériences seront menées pour rechercher les bonnes configurations des systèmes proposés.
- Finalement, une conclusion générale résume les différents travaux effectués ainsi que les perspectives permettront d'élargir et de poursuivre l'étude menée dans cette thèse de Doctorat.

CHAPITRE I: GENERALITES SUR L'IDENTIFICATION DES APPARELS ELECTRIQUES

I-1 introduction

L'évolution progressive croissante de la consommation de l'énergie domestique mondiale a poussé les fournisseurs d'énergie électrique à vivre en permanence dans un esprit de concurrence rehaussant ainsi leurs capacités à innover pour optimiser la consommation et améliorer les performances de services en termes de quantité et surtout en qualité. Dans cette optique, la fonction des compteurs traditionnels qui consistait à fournir uniquement des informations sur la quantité d'énergie consommée, a été développée par l'invention de compteurs intelligents (smart meter » en anglais) conçus spécialement pour donner des informations sur l'identité de l'appareil, le temps et la quantité de puissance consommée aux distributeurs d'électricité. Cette fonction est en effet réalisée, en intégrant dans le processus de traitement de ces compteurs, une phase d'identification des appareils électriques en temps réel permettant aux clients de maîtriser leurs consommations électriques et aux fournisseurs d'énergie de mieux gérer leurs distributions énergétiques. Le processus d'identification des appareils électriques est mis en place en fonction des spécifications des appareils telles que les types d'appareils préférés par les clients (appareils les plus utilisés par les clients), leurs modes de fonctionnement (appareils à deux états, à états multiples, à variations continue et permanente), les caractéristiques de leurs formes d'onde (du courant ou de la tension) et leurs interfaces utilisateurs (appareils activés ou auto-activés par l'utilisateur) [[HYPERLINK \I "WWi15" 8 \],9](#)}, [[HYPERLINK \I "SSH17" 10](#)]. Dans ce chapitre, on présente des généralités sur l'identification des appareils électriques, ainsi que les différentes méthodes de classification utilisées.

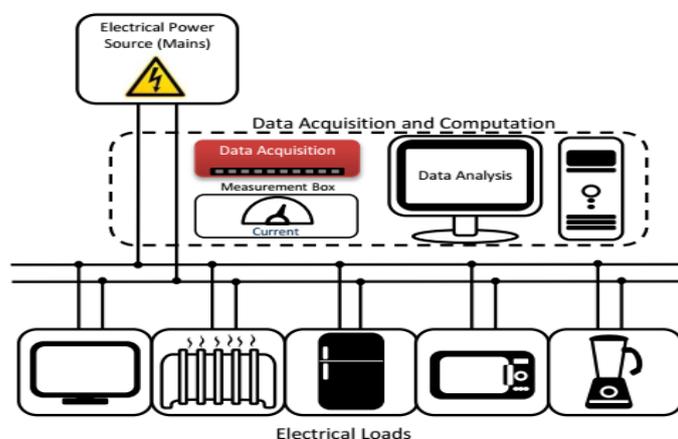


Figure I-1 Composants principaux d'un système de surveillance de la charge des appareils électriques [11](#)}}

I-2 Objectifs d'identification des appareils électriques

Il existe plusieurs approches dans lesquelles l'identification des appareils électriques est mise en œuvre dans un environnement domestique [[HYPERLINK \l "MZe11" 12](#)], [13](#)]. L'identification des appareils électriques, dans leur état le plus simple, est utilisée comme méthode de réduction de la consommation d'énergie. Actuellement, les fournisseurs d'électricité s'efforcent de réduire leur utilisation carbonique et de combler l'écart énergétique entre la production d'électricité durable et la consommation d'énergie. L'utilisation de la surveillance de l'énergie et des compteurs intelligents peuvent aider à réduire cette consommation électrique domestique. Les connaissances sur les principaux appareils consommateurs d'énergie, obtenues par la tâche d'identification, permet aux clients d'ajuster leurs comportements en réduisant une partie de leurs consommations d'énergie.

Du point de vue réseau intelligent, la gestion de la demande pourrait être complétée par une analyse de la consommation d'énergie aux heures de pointe. Un exemple de gestion de la demande est d'encourager les consommateurs à réduire leur consommation d'énergie pendant ces heures et de déplacer cette utilisation aux heures creuses. Cela donnera une image plus approfondie de la consommation d'énergie et éventuellement une indication de l'endroit où la consommation d'énergie peut être réduite. Les fournisseurs de services publics estiment que l'investissement dans le réseau intelligent, qui comprend des compteurs intelligents, leur fournira des capacités croissantes pour la gestion, au fil du temps, de demande de l'énergie. Dans le cadre de ces nouvelles capacités, la communication et la gestion des données jouent un rôle important. Celles-ci entraînent potentiellement des améliorations dans des domaines tels que la réponse à la demande, la capacité de se connecter au réseau avec des micro-réseaux, de brancher des véhicules électriques et de stocker l'énergie [[HYPERLINK \l "HFa10" 14](#)].

Une autre utilisation de l'identification des appareils électriques -non étudiée- consiste à combiner l'identification des appareils avec la surveillance de l'état du signal électrique et la détection des défauts. De nombreux travaux, basés sur la transformée de Fourier rapide (FFT: Fast Fourier Transform) ou l'ondelette (DWT: Discret Wavelet Transform), ont été menés par l'analyse des signaux électriques en vue d'identifier l'état

des pompes industrielles et détecter leur défauts 15]], [[HYPERLINK \I "NMe10" 16](#)]. Ces travaux démontrent que la FFT du signal électrique de la pompe induit des pics supplémentaires dans le spectre quand il y a un défaut. En effet, pour chaque défaut il y a un pic différent dans le spectre permettant ainsi l'identification de multiples défauts. Dans 17]], il a été proposé que le signal transitoire d'une charge puisse indiquer si un dispositif est détérioré. Cette dernière peut techniquement être détectée à l'aide d'une identification de charge en fonction de la résolution du système et le degré du défaut. Un exemple d'utilisation d'une combinaison d'identification de charge et de surveillance de l'état électrique peut être vu dans la référence [[HYPERLINK \I "WCG05" 18](#)]. Ceci offre la possibilité à un système d'identification de charge d'identifier les moments où les appareils fonctionnent en dehors de leur comportement normal et les moments où ces appareils fonctionnent de manière inefficace. La mise en œuvre de l'identification des appareils avec surveillance de l'état pourrait hypothétiquement permettre une identification précoce des appareils défectueux, empêchant subséquemment les pannes inattendues.

Un autre exemple où l'identification de l'appareil électrique peut être utilisée, est la détection de l'activité à la maison. Cela a une gamme d'applications, y compris les soins de santé, le divertissement, la domotique et la surveillance de l'énergie.

Généralement, les méthodes d'identification des appareils électriques se regroupent en deux approches. La première, nommée ILM (Intrusive Load Monitoring), est appliquée à l'intérieur de l'habitat et la deuxième, nommée l'approche NILM (Non-Intrusive Load Monitoring) fonctionne au niveau du compteur principal 19]]. Dans notre travail, on s'intéresse à l'approche NILM.

I-3 Identification des appareils électriques dans l'approche NILM

L'approche NILM est l'une des méthodes de contrôle et de surveillance de la consommation énergétique domestique et industrielle par la mise en place d'un seul compteur à l'entrée de l'habitat.

Le cadre général de l'approche NILM commence à partir des mesures faites sur l'entrée de la consommation totale d'électricité pour finir finalement parla décomposition, en contributions individuelles, de chaque appareil. Du point de vue

économique, l'approche NILM est la plus adaptée puisqu'elle n'utilise qu'un seul point de mesure au niveau du disjoncteur principal. A partir du signal mesuré, on peut extraire des informations sur l'identité de l'appareil en état de fonctionnement ainsi que d'autres informations telles que l'intervalle et la quantité de puissance consommée. La Figure I-2 présente schéma-bloc des différentes méthodes de surveillance des appareils électriques domestiques.

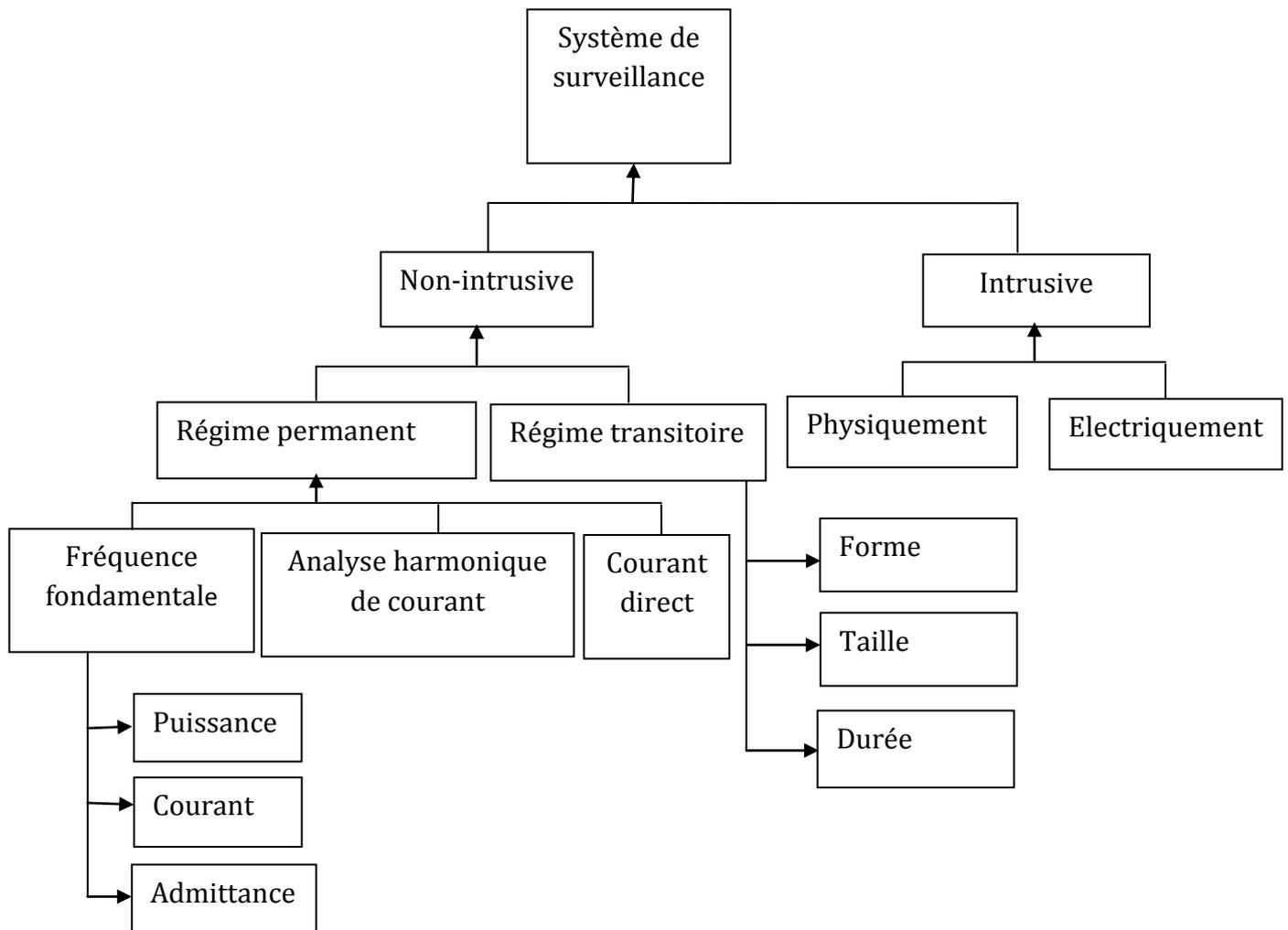


Figure I-2 Classification des méthodes de surveillance de la consommation d'énergie électrique [[HYPERLINK \I "VHa14" 19](#)]

Dans la référence [20](#)}, [[HYPERLINK \I "Har92" 21](#)], Hart a présenté les premiers travaux sur les méthodes NILM. Il a été le premier à analyser les variations de la puissance totale pour identifier les appareils électriques, mais la méthode ne permet pas d'identifier certains appareils (appareils multi-états comme les machines à laver). Les travaux de base de Hart ont lancé plusieurs travaux de recherches notamment ceux de

Sultanem 22}} qui a utilisé les variations de courant et les puissances active et réactive comme paramètres pertinents pour l'identification. Les Modèles de Markov Cachés (HMM: Hidden Markov Models) ont d'abord été utilisés dans [HYPERLINK \l "Bon96" 23]. Dans la référence 24}}, les auteurs ont utilisé des harmoniques de courant afin d'identifier la nature structurelle des appareils électriques. Leeb, dans [HYPERLINK \l "Lee93" 25], a proposé les premiers travaux qui se sont intéressés à la partie transitoire du courant. Il a étudié les formes transitoires des puissances active et réactive. Beaucoup d'autres travaux comme ceux présentés dans les références 26}}, [HYPERLINK \l "Dre99" 27], 28}}, [HYPERLINK \l "Bar03" 29], 2}}, [HYPERLINK \l "Cha08" 30], 31}}, [HYPERLINK \l "Car13" 32] avec un état de l'art présenté dans 33}}, se sont intéressés aux NILM pour l'identification des appareils électriques domestiques.

Comme on peut le constater, à partir des travaux se focalisant l'approche NILM, le point critique dans cette dernière est la phase d'extraction des paramètres descripteurs de l'appareil à partir de la tension $u(t)$ et du courant $i(t)$ mesurés à l'entrée. L'objectif est donc l'extraction, pour chaque appareil, d'une signature sous forme d'une combinaison de plusieurs paramètres qui peuvent être de différentes natures plus particulièrement des puissances actives ou réactives, des harmoniques de courant ...etc.

I-4 Principe de fonctionnement d'un système IAED

L'identification des appareils électriques domestiques est le processus d'analyse des signaux émis par ces derniers pour identifier l'appareil en état de fonctionnement. Comme nous l'avons déjà discuté, les signaux classifiables émis par les appareils peuvent être le courant électrique, la tension électrique et la puissance consommée. Ces signaux sont analysés pour les caractériser et classifier par conséquent les différentes classes d'appareils.

Un système IAED fonctionne sur plusieurs étapes. La première étape consiste à détecter l'événement de la mise en marche de l'appareil à identifier. Cette détection est appliquée pour décider si un appareil est connecté ou déconnecté au système. La deuxième étape, utilisée dans le cas où un appareil est connecté, comprend l'analyse des signaux par un module d'extraction des paramètres caractérisant les différentes classes d'appareils. La troisième étape est l'identification des appareils électriques, qui fait

correspondre le signal électrique d'entrée du système à une classe d'appareils en utilisant l'un des classificateurs standards tels que le KNN (K Nearest Neighbour), le SVM (Support Vector Machine), ...etc. La Figure I-3 présente le schéma synoptique des différentes étapes du processus d'identification IEAD.

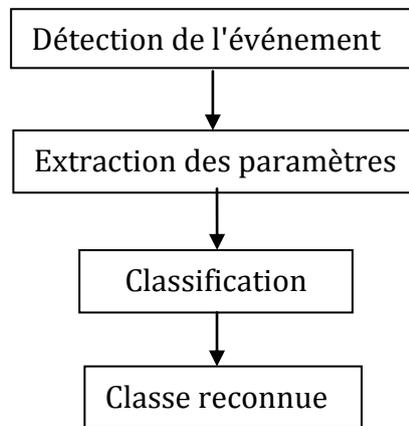


Figure I-3 Schéma synoptique général du processus d'identification IAED

La conception d'un IAED est basée sur deux phases principales. La première est la phase d'apprentissage pour modéliser les classes d'appareils, quant à la deuxième c'est une phase d'identification utilisée pour classifier les signaux de test et évaluer les performances du système. Les différentes phases de conception d'un IAED et leurs étapes sont illustrées sur la Figure I-4. Ces deux phases exigent une étape d'extraction des paramètres représentant les différentes classes d'appareils, ainsi qu'une étape de classification des signaux de test à partir des modèles entraînés durant la phase d'apprentissage.

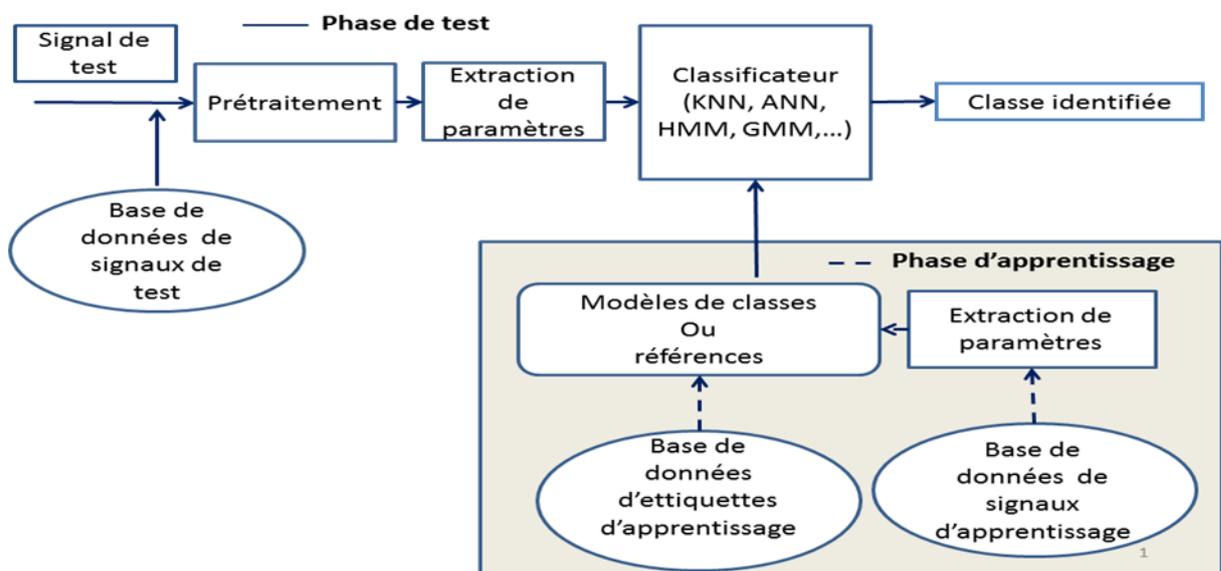


Figure I-4 Schéma synoptique de la conception d'un système IAED

La phase d'apprentissage consiste à entraîner les modèles de classes d'appareils ou de stocker des références sous forme de vecteurs de paramètres dans le but de les exploiter dans l'étape de classification. La phase de test consiste à évaluer les performances, en termes de précision et de complexité, du système. Généralement, la précision du système est évaluée en utilisant le taux de classification (CR) et la complexité est évaluée en utilisant le temps de calcul ainsi que l'espace mémoire.

L'évaluation des performances du système exige une base de données de référence répartie en une base d'apprentissage et une base de test. Pratiquement, les différentes étapes de conception d'un système IAED sont données comme suit :

- Préparation de données ;
- Modélisation des différents modèles de classes en utilisant une base de données d'apprentissage ;
- Classification des signaux d'entrée du système ;
- Evaluations des performances sous une base de données de test ;

Dans les sections suivantes, nous allons présenter un état de l'art sur les techniques de classification des signaux ainsi que les différentes mesures d'évaluation de performances des systèmes IAED. Les techniques d'extraction de paramètres seront détaillées dans le chapitre II.

I-5 Etat de l'art sur les systèmes d'identification des appareils électriques

Au cours des vingt dernières années, de nombreux systèmes ont été proposés pour surveiller et identifier les appareils électriques domestiques. Tout système IAED se base sur une étape d'extraction des paramètres à partir du signal électrique et une étape de classification basée sur un algorithme de classification tels que l'algorithme HMM [[HYPERLINK \l "Rid14" 34](#)], [1](#)], celui de SVM [[HYPERLINK \l "Mit13" 35](#)], [2](#)], [[HYPERLINK \l "MBe10" 36](#)], l'algorithme réseaux de neurones artificiels (ANN: Artificial Neural Network) [37](#)], [[HYPERLINK \l "DSr06" 38](#)] et celui des KNN [39](#)].

Dans [[HYPERLINK \l "Rid14" 34](#)], les auteurs ont appliqué l'algorithme de classification HMM sur des signaux de la base de données ACS-F1 échantillonnée à 10 Hz (basse fréquence).

Dans [1](#)], les auteurs ont appliqué également l'algorithme HMM combiné avec une méthode d'extraction basée sur l'analyse harmonique du signal de courant électrique. Les performances de leur système IAED sont évaluées sous la base de données PLAID échantillonnée à 30KHz. Dans [[HYPERLINK \l "Hac181" 40](#)], les auteurs ont amélioré les performances du système proposé dans [1](#)] en se basant sur l'analyse par ondelettes.

Dans [[HYPERLINK \l "Pat07" 2](#)], les auteurs ont utilisé le classificateur SVM combiné avec l'extraction des paramètres du régime transitoire du courant électrique. Dans [36](#)], les auteurs ont utilisé des paramètres extraits des puissances réelle et réactive.

Dans [[HYPERLINK \l "HHC12" 37](#)], [38](#)], les auteurs ont utilisé le classificateur ANN combiné avec l'extraction des paramètres à partir des puissances réelle et réactive, ainsi que des paramètres des événements transitoires. Plus particulièrement, dans [[HYPERLINK \l "DSr06" 38](#)], les auteurs ont proposé d'utiliser les quinze premières harmoniques du courant.

Dans [4](#)], les auteurs ont proposé un système IAED basé sur le classificateur KNN combiné avec l'extraction d'un seul vecteur de paramètres représentant entièrement le signal du courant électrique.

Dans notre travail on s'intéresse à ce type de classificateur vu sa simplicité et son efficacité. Le tableau I-1 présente un état de l'art récapitulatif sur les techniques de classification utilisées dans les systèmes IAED.

Tableau I-1: état de l'art récapitulatif sur les techniques de classification utilisées dans systèmes IAED.

Références	Le classificateur utilisé	La base de données utilisée	Type du signal électrique
[HYPERLINK \l "Rid14"]	HMM/GMM	ACS-F1	Courant/tension électriques

Extraction et sélection des paramètres pertinents appliquées à l'identification des appareils électriques

34]			
1]]	HMM	PLAID	Courant électrique
[HYPERLINK \\ "Hac181" 40]	HMM	PLAID	Courant électrique
35]]	SVM	/	Puissance active
[HYPERLINK \\ "Pat07" 2]	SVM	/	Bruit da la tension électrique
36]]	SVM	/	Puissance réels et réactifs.
[HYPERLINK \\ "HHC12" 37]	ANN	/	Puissance électrique
38]]	ANN	/	Courant électrique
[HYPERLINK \\ "Nai171" 4]	KNN	COOLL	Courant électrique

I-6 Algorithmes de classification

L'étape de classification est une étape très essentielle dans le processus de conception d'un système IAED. Cette étape peut s'effectuer de deux manières soit en classifiant une séquence de vecteurs de paramètres représentant le signal d'entrée 40]], [HYPERLINK \\ "Nai161" 1], ou en classifiant un vecteur de paramètres représentant entièrement le signal d'entrée 4]]. Plus particulièrement, dans [HYPERLINK \\ "Nai161" 1], les auteurs ont appliqué la première approche en classifiant chaque séquence de vecteurs de paramètres harmoniques extraite d'un signal de courant électrique de la base de données de test en une classe d'appareils par l'algorithme de classification HMM.

Cependant, dans [4]}, les auteurs ont utilisé la deuxième approche en appliquant l'algorithme de classification KNN sur un seul vecteur de paramètres extraits à partir d'un modèle paramétrique représentant le signal du courant électrique.

Dans notre travail de thèse, la tâche de classification est basée sur une approche hybride qui consiste à classifier initialement chaque vecteur de paramètres d'une séquence de vecteurs en une classe d'appareils. Ensuite, la stratégie de règle de vote sera appliquée sur la séquence d'indices de classes pour décider sur la classe du signal d'entrée. Cette approche sera détaillée dans le chapitre IV.

Dans les sous-sections suivantes nous allons décrire les principaux classificateurs supervisés utilisés pour la tâche IAED. Ces classificateurs exigent une phase d'apprentissage des différentes classes d'appareils en utilisant une base d'apprentissage.

I-6-1-Classificateur Machines à Vecteurs de Support SVM

La bonne réputation de la méthode SVM, pour la classification binaire en particulier, provient du fait qu'elle repose sur l'application d'algorithmes de recherche de règles de décision linéaires ("hyper plan séparateur"). Néanmoins, la recherche s'effectue dans un espace ("feature space") de très grande dimension. Les machines à vecteurs de support (SVM) ont été utilisées dans plusieurs systèmes d'identification des appareils électriques. En effet, Patel dans [HYPERLINK \l "Pat07" 2], a utilisé des informations du courant transitoire à haute fréquence et un SVM pour identifier les appareils électriques. Cette méthode de classification a été également utilisée dans [36]} avec des paramètres extraits des signaux de puissance réels et réactifs. La classification est effectuée géométriquement, ce qui signifie que chaque classe occupe un espace dérivé des données de formation. La meilleure fonction de classification est trouvée en maximisant la marge entre les classes. Le premier inconvénient des méthodes SVM est qu'elles sont inefficaces lorsque le nombre de classes devient élevé. Cependant, pour les algorithmes d'identification des appareils électriques, cet inconvénient ne représente pas une limitation du fait que le nombre maximum d'appareils avoisine quarante [HYPERLINK \l "TO00" 41].

L'algorithme SVM, décrit dans [41]}, est un outil puissant spécifique pour la classification des données. La première étape majeure d'une classification SVM est de

construire un plan de décision qui sépare un ensemble d'objets avec différentes appartenances de classes. Il garantit la meilleure fonction permettant de distinguer les membres des classes en maximisant la marge entre eux. Les hyper-plans de marge maximale permettent les meilleures capacités de généralisation et donc les meilleures performances de classification sur l'ensemble de données d'apprentissage. Cette procédure nécessite la recherche de la solution du problème d'optimisation suivant:

$$\min_{W, b, \varepsilon} \left(\frac{1}{2} W^T + C \sum_{i=1}^l \varepsilon_i \right) \quad (I-1)$$

$$\text{sachant que } \begin{cases} y_i (W^T \phi(x_i) + b) \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0 \end{cases} \quad (I-2)$$

Avec :

l : le nombre total de sous-séquences,

W : le vecteur normal de l'hyper-plan,

b est le décalage de l'hyper-plan,

C est le paramètre de pénalité du terme d'erreur ξ et $\Phi(X)$ est la fonction noyau de projections .

La deuxième étape majeure d'une classification SVM consiste à choisir la fonction noyau de l'algorithme. En effet, pour deux groupes i et j , les vecteurs d'apprentissage x_i et x_j sont transformés vers un espace de dimension supérieure en utilisant la fonction de noyau φ définie comme suit :

$$K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j) \quad (I-3)$$

La méthode nécessite donc de sélectionner un noyau K ainsi que d'autres paramètres. Parmi les choix possibles, on compte en particulier :

Le noyau linéaire donné par :

$$K(x, x') = \langle x, x' \rangle \quad (I-4)$$

Le noyau Gaussien radial donné par :

$$K(x, x') = \exp(-\sigma \|x - x'\|^2) \quad (I-5)$$

Les noyaux polynomiaux donnés par :

$$K(x, x') = (\alpha + \beta \langle x, x' \rangle)^\delta \quad (I-6)$$

Le noyau radial de Laplace donné par :

$$K(x, x') = \exp(-\sigma \|x - x'\|) \quad (I-7)$$

Le noyau tangente hyperbolique (sigmoïde) donné par :

$$K(x, x') = \tanh(\alpha + \beta \langle x, x' \rangle) \quad (I-8)$$

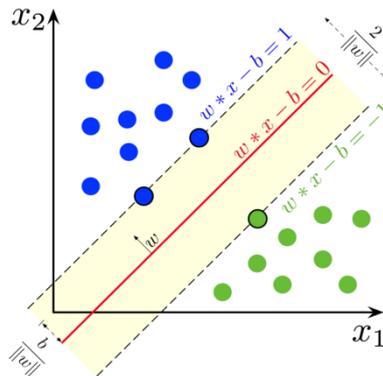


Figure I-5 Hyper-plan à marge maximale pour un SVM formé avec des échantillons de deux classes [[HYPERLINK \l "TOu00" 41](#)]

I-6-2 Classificateur K-plus proche voisin (KNN: K-Nearest Neighbours classifier)

Le classificateur KNN a été utilisé par [42](#)] pour classer les appareils utilisant des informations de tension à haute fréquence. En effet, il cherche, dans l'ensemble d'apprentissage, un groupe de K objets les plus proches de l'objet de test et attribue l'étiquette en fonction du voisinage par l'application de la stratégie de vote majoritaire [[HYPERLINK \l "XWu08" 43](#)]. Cette approche est basée sur un ensemble d'objets étiquetés en utilisant une mesure de similarité pour calculer la distance entre un objet de test et les K plus proches voisins. Il y a un certain nombre de paramètres qui doivent être fixés avant d'implémenter le classificateur KNN, par exemple le choix de la valeur K, et

également la méthode de mesure de la distance entre les objets. Généralement, le classificateur KNN est appliqué en choisissant la distance euclidienne avec K égal à 1.

Dans les systèmes IAED, c'est le classificateur KNN qui est couramment utilisé vue sa simplicité d'implémentation. Réellement, pour toute nouvelle instance (objet) de données, les attributs du nouveau cas (test) sont comparés à tous les cas précédemment vus dans la base de données d'apprentissage. La comparaison est généralement basée sur des mesures de distance. Les instances ou les cas les plus proches dans la base de données d'apprentissage sont évalués en fonction de la mesure de distance. La nouvelle instance est affectée à la classe de la majorité des instances voisines (processus de classification). Mathématiquement parlant, pour toute instance x_i dans la base de donnée de test, la distance minimale $d(x_i)$ consiste à chercher l'instance x_j appartenant de la base d'apprentissage, et qui permet de minimiser la distance $d(x_i, x_j)$:

$$d(x_i) = \min_j d(x_i, x_j) \quad j \in \{1, 2, \dots, n\} \quad (I-9)$$

I-6-2-1 Quelques exemples de distances

Les distances couramment utilisées sont la distance euclidienne, la distance corrélation, la distance Cityblock, la distance cosinus et la distance de Hamming.

Dans le but de décrire ces distances, on suppose deux vecteurs X et Y de dimension d donnés par: $X = [x_1, x_2, \dots, x_d]$, et $Y = [y_1, y_2, \dots, y_d]$.

Dans ce qui suit seront définies quelques exemples de distances.

I-6-2-1-a Distance euclidienne

La distance euclidienne est une métrique très utilisée dans la classification KNN. Elle est définie par la formule suivante:

$$d_E(X, Y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2} \quad (I-10)$$

I-6-2-1-b Distance de Cityblock

La distance Cityblock ($D_{\text{cityblock}}$) est définie par la somme des différences absolues entre les composantes (x_i, y_i) . Elle est donnée par :

$$D_{\text{cityblock}}(X, Y) = \sum_{i=1}^d |x_i - y_i| \quad (\text{II} - 11)$$

I-6-2-1-c Distance Cosinus

La distance cosinus (D_{cos}), appelée souvent la distance angulaire, est calculée en soustrayant la valeur de similarité cosinus de la valeur 1. Cette distance est dérivée de la mesure de similarité entre deux vecteurs [44]. Elle est définie comme suit:

$$D_{\text{cos}}(X, Y) = 1 - \frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \quad (\text{I} - 12)$$

I-6-2-1-d Distance Corrélation

La distance corrélation, dérivée de la distance de Pearson, est calculée à partir du coefficient de corrélation de Pearson. Ce dernier mesure la relation linéaire entre deux vecteurs. Cette distance est obtenue en soustrayant la valeur du coefficient de corrélation de Pearson de la valeur 1. Elle est normalisée par rapport à sa valeur maximale pour obtenir une mesure de distance $D_{\text{corr}}(X, Y)$ comprise entre 0 et 1. Elle est donnée par :

$$D_{\text{corr}}(X, Y) = \frac{1}{2} \left(1 - \frac{\sum_{i=1}^d (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^d (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^d (y_i - \bar{y})^2}} \right) \quad (\text{I} - 13)$$

$$\text{où : } \bar{x} = \frac{1}{d} \sum_{i=1}^d x_i \text{ et } \bar{y} = \frac{1}{d} \sum_{i=1}^d y_i$$

La valeur $D_{\text{corr}}(X, Y) = 1$ indique un comportement similaire et la valeur $D_{\text{corr}}(X, Y) = 0$ un comportement opposé entre X et Y.

I-6-2-1-e Distance de Hamming

La distance de Hamming est une métrique qui mesure le nombre de discordances (la différence) entre deux vecteurs. Elle définit comme suit:

$$D_{Ham}(X,Y) = \sum_{i=1}^L 1_{x_i \neq y_i} \quad (I - 14)$$

I-6-2-2L'algorithme de base du classificateur KNN

L'algorithme de base du classificateur KNN est donné comme suit :

Début d'algorithme

- Introduire les entrées ;

- vecteurs de paramètres d'apprentissage VAPP,
- vecteurs de paramètres de test VT,
- étiquettes de classes des vecteurs VAPP,
- la valeur de K,
- la distance utilisée (choisie).

- Pour chaque vecteur VT_i de (VT):

1. calculer les distances entre chaque vecteur de la base de test VT_i et tous les vecteurs de la base d'apprentissage VAPP,
2. Chercher les K vecteurs de la base d'apprentissage VAPP qui sont proches, au sens de la distance, du vecteur VT_i,
3. Attribuer au vecteur VT_i la classe la plus dominante ou la plus votée,
4. Refaire les étapes 1, 2 et 3 pour tous les vecteurs de la base de test VT.

Fin de l'algorithme

I-6-1-4 Modèles de Markov cachés (HMM: Hidden Markov Models)

Un modèle de Markov est une technique de modélisation statistique probabiliste d'états finis, largement utilisée dans le domaine du traitement du signal. Un HMM est décrit par un ensemble de nœuds (ou états) reliés entre eux par des arcs de transitions. Cette technique est contrôlée par deux processus stochastiques. Le premier commence sur l'état initial du HMM et se déplace ensuite d'un état à un autre à chaque instant t ($1 \leq t \leq T$), en respectant les transitions autorisées par la topologie. Le deuxième processus génère, après chaque changement d'état à l'instant t , une observation " o_t ". À

chaque état i ($1 \leq i \leq N$) est associée une distribution de probabilité $b_i(o)$ et à chaque transition de l'état i à l'état j est associée une probabilité de transition a_{ij} . La distribution de probabilité $b_i(o)$ représente la probabilité d'émission sur l'état i de l'observation o [45]. La Figure I-6 présente un exemple d'un Modèle de Markov gauche-droite à 4 états.

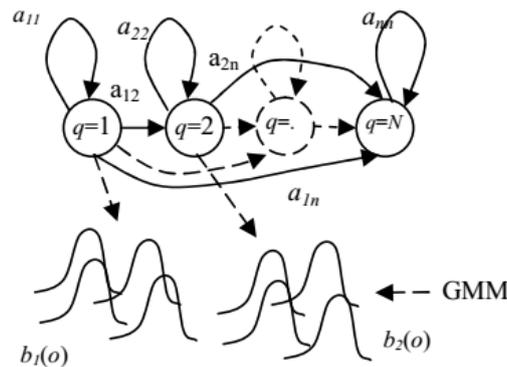


Figure I-6 Exemple d'un Modèle de Markov gauche-droite [45]

Les modèles de Markov cachés (HMM) sont largement utilisés dans le domaine d'identification des appareils électriques [46], [47], [48], [49]. Ils sont considérés comme un classificateur utilisé avec des mesures à basse fréquence, dans l'identification des appareils électriques. Dans ce cas, les données peuvent être la puissance réelle ou le courant électrique [50]. Le choix d'un algorithme HMM est populaire car les modèles sont mathématiquement riches et lorsqu'ils sont appliqués correctement, ils fonctionnent bien dans la pratique. Il y a quatre composants principaux dans un HMM. Ces derniers sont les états où les étiquettes sont attribuées, les probabilités d'émission (chaque état a sa propre probabilité d'émission qui est basée sur les paramètres de chaque classe), les probabilités de transition (probabilité de passer d'un état à un autre) et la probabilité de sortie [51]. Chaque appareil est modélisé comme un HMM unique formé à l'aide d'un certain nombre d'observations. Des exemples d'observations utilisées pourraient être la probabilité initiale d'un état d'appareil, le nombre d'états possibles qu'un appareil peut avoir ou la probabilité que l'appareil soit allumé à un moment donné. Le système utilise une observation pour déduire quel état a changé. Ici, la tâche consiste à identifier les appareils électriques compte tenu des paramètres du modèle. La séquence d'observation est utilisée pour ajuster les paramètres du modèle

pendant la séquence d'apprentissage. Le problème de l'apprentissage est capital pour la plupart des applications des HMM, car il permet d'adapter les paramètres du modèle aux séquences observées [11].

I-7. Mesures des performances d'un système d'identification

Les performances d'un système de classification sont évaluées en utilisant des mesures telles que le taux de classification (classification ratio: CR), la précision p, la sensibilité S, le F-score et la matrice de confusion. Dans la sous-section suivante, nous allons présenter les différentes mesures d'évaluation des performances couramment utilisées dans les systèmes IAED [12], [52], [53], [54].

I-7.1 Taux de Classification TC (RC : classification ratio)

Le RC est l'une des mesures, utilisée pour l'évaluation des performances des classificateurs, le plus utilisées. Elle donne une évaluation globale de l'opération de classification. Cette mesure s'effectue en comparant les résultats de sortie du classificateur avec les classes originales des signaux de la base de test. Elle est définie par:

$$CR(\%) = \frac{T - M}{T} * 100 \quad (I - 15)$$

où T est le nombre total des signaux testés et M est le nombre des signaux testés mal classés.

I-7-2 Matrice de confusion

Une matrice de confusion est une disposition de table spécifique qui permet de visualiser les performances détaillées d'un algorithme de classification. Cette matrice contient des informations sur les résultats réels et prévus obtenus par un système de classification [55]. Le tableau I-2 montre la matrice de confusion pour un classificateur dont chaque colonne de la matrice représente les instances d'une classe prédite, tandis que chaque ligne représente les instances d'une classe réelle. Le nom de cette métrique d'évaluation des performances vient du fait que l'on peut voir si le système confond deux classes. Les meilleures performances d'une matrice de confusion auront 100% le long de la diagonale.

Table I-2 Description de la matrice de confusion dans la classification binaire [56]

	classe prédite	
	1	0
Déecté	1	TP
Non déecté	0	FN

Dans ce tableau, nous avons :

TP est le nombre de vrais positifs, c'est-à-dire les échantillons positifs correctement classés ;

FP est le nombre de faux positifs, c'est-à-dire les échantillons négatifs incorrectement classés ;

FN est le nombre de faux négatifs, c'est-à-dire les échantillons positifs incorrectement classés ;

TN est le nombre de vrais négatifs, c'est-à-dire les échantillons négatifs correctement classés.

I-7-3 Mesures statistiques

La matrice de confusion permet de définir plusieurs mesures statistiques telles que la précision, la sensibilité et le F-scor. Ces mesures sont calculées à partir des valeurs de TP, FP, TN et FN. Elles sont données comme suit:

- ✓ **Précision** : C'est le pourcentage d'échantillons positifs correctement classés sur le total des échantillons classés positivement. Ce pourcentage est donné par :

$$P(\%) = \frac{TP}{TP + FP} \cdot 100 \quad (I - 16)$$

- ✓ **Sensibilité** : C'est le pourcentage d'échantillons positifs correctement classés sur le total des échantillons positifs. Ce pourcentage est donné par :

$$S(\%) = \frac{TP}{TP + FN} \cdot 100 \quad (I - 17)$$

✓ **F-score:** C'est une moyenne pondérée de la précision et de la sensibilité. Elle est donnée par :

$$F - scor(\%) = \frac{2 P S}{P + S} = 2 \cdot \frac{TP}{2TP + FP + FN} \quad (I - 18)$$

Ces mesures sont évaluées pour chaque classe d'appareils électriques. Donc, ces mesures sont considérées comme des mesures d'évaluation locale [12], [52], [53].

I-8 Approche de classification proposée pour la conception d'un système IAED

L'identification d'un appareil à partir d'un signal électrique, exige deux étapes. En effet, la première est une étape d'extraction de paramètres permettant de convertir le signal en une séquence de vecteurs de paramètres. La deuxième est une étape de classification de la séquence résultante en une classe d'appareils en utilisant un classificateur. Dans [1], les auteurs ont proposé l'application du classificateur HMM sur la séquence de vecteurs de paramètres harmoniques STFS représentant le signal. Dans [4], les auteurs ont proposé l'application du classificateur KNN sur un vecteur de paramètres représentant le signal d'entrée. Dans notre travail, on s'intéresse à l'application du classificateur KNN sur chacun des vecteurs de paramètres représentant le signal d'entrée appartenant à la même séquence. Cette classification exige une étape de prise de décision sur la classe d'appareils appartenant à cette séquence. Dans notre travail, nous proposons d'appliquer la stratégie de la règle de vote pour décider sur la classe d'appareils [5] [6] [7]. Dans la section suivante, on présentera un aperçu sur cette stratégie et ses applications.

I-8-1 Stratégie de règle de vote

La Stratégie de Règle de Vote (SRV) est l'une des techniques les plus efficaces dans le domaine de la reconnaissance de formes. Dans cette stratégie, la décision de classification finale est la décision à la majorité prise parmi un ensemble de nombreux systèmes concurrents [57], [58]. Dans [57], les auteurs ont donné un aperçu et une comparaison des méthodes de classification basées sur la SRV appliquée également dans l'identification des appareils électriques par la combinaison de plusieurs classificateurs [58]. Dans notre travail contrairement à l'application de la SRV entre plusieurs classificateurs [58], nous proposons l'application de la SRV sur une séquence d'indices

de classes. Cette séquence est en fait obtenue par la classification KNN de chaque vecteur d'une séquence de vecteurs de paramètres statistiques, pour identifier la classe d'appareil la plus votée. L'application de cette stratégie pour la tâche IAED sera détaillée dans le quatrième chapitre.

I-9 conclusion

Dans ce chapitre, nous avons présenté un état de l'art qui nous a permis de faire le point sur les différentes techniques d'identification des appareils électriques domestiques (IAED) à partir de leurs consommations énergétiques globales. Plus particulièrement, la méthode d'identification de type NILM ainsi que son principe de fonctionnement ont été brièvement décrits. Plusieurs algorithmes de classification, utilisés dans les systèmes IAED, ont été introduits de la conception d'un système IAED.

À ce jour, plusieurs méthodes d'identification des appareils électriques ont été proposées pour obtenir de bonnes performances en termes de précision et de complexité. Concernant l'étude des performances, nous avons montré que le taux de classification n'est pas la seule mesure sur laquelle l'efficacité d'une technique d'identification des appareils électriques est mesurée. Pour cette raison, différentes mesures ont été définies à savoir la matrice de confusion, la précision(P), la sensibilité(S), et le F-score.

CHAPITRE II:
METHODES D'EXTRACTION DE
PARAMETRES

II-1 Introduction

Ce chapitre a pour objectif de décrire les différentes méthodes d'extraction des paramètres et leurs applications dans les systèmes IAED. La plupart des méthodes d'extraction, utilisées pour la tâche IAED, se basent sur les techniques de transformées des signaux telles que la transformée de Fourier [1], la transformée par ondelettes [40], ...etc.

La section suivante donnera un état de l'art sur les différents descripteurs (type de paramètres) et techniques, utilisés pour la tâche IAED.

II-2 Etat de l'art sur l'extraction des paramètres

L'étape d'extraction des paramètres par les systèmes IAED est une étape indispensable pour la tâche d'identification des appareils domestiques à partir des mesures électriques telles que les mesures du courant électrique, de la tension et de la puissance instantanée. Plusieurs bases de données, utilisées dans la conception des systèmes IAED, contiennent des mesures du courant et de la tension électriques. Généralement, le signal du courant électrique est utilisé comme un signal d'entrée dans la phase d'extraction des paramètres pour la tâche IAED. Cette phase a pour objectif de réduire le flux de données en extrayant les paramètres discriminants des différentes classes d'appareils électriques. Elle consiste généralement à découper le signal d'entrée en une séquence de fenêtres chevauchées dont chacune sera transformée en un vecteur de paramètres. Ainsi, le signal d'entrée est converti en une séquence de vecteurs de paramètres.

Plusieurs travaux de recherches approfondies sur l'extraction des paramètres discriminants dans l'approche NILM sont inspirés du travail fondateur de Hart et al. [20], [21]. Zeifman et al. [12] et Zoha et al. [59] ont fourni un aperçu complet sur les algorithmes d'extraction des paramètres qui peuvent être extraits à partir des mesures de tension ($V(t)$) ou de courant ($I(t)$). Ces paramètres et leur nombre dépendent de la fréquence d'échantillonnage [12] et sont classés en paramètres d'états stable et transitoire [59]. Dans les sections suivantes, nous allons présenter les différents types de paramètres selon la nature du signal électrique utilisé pour l'extraction des paramètres discriminants.

II-2-1 Puissance réelle

Bien que la mesure de la puissance réelle soit simple à obtenir en comparaison avec celle de la puissance réactive, les appareils, ayant des consommations d'énergie similaires, sont plus difficiles à discriminer en s'appuyant uniquement sur des mesures de puissance. Par conséquent, ce paramètre de puissance a été généralement complété par le temps et la fréquence d'utilisation des appareils. Dans [60], Powers et al ont enregistré le temps et la fréquence de l'occurrence des grands changements dans les mesures de puissance à partir d'un ensemble de données échantillonnées (avec une durée de 15 min). Dans la référence [61], Farinaccio et al ont utilisé des règles de décision spécifiques à l'appareil, tandis que Marceau et al., dans la référence [62] ont utilisé le filtrage du signal (lissage) ainsi que les statistiques de durée d'utilisation comme informations complémentaires. Enfin, Baranski et al. [29] ont complété les données de puissance réelle avec un histogramme de fréquence.

II-2-2 Paramètres PQ

Les changements progressifs de la puissance active (P) et de la puissance réactive (Q) sont les premières signatures couramment utilisées en régime permanent dans les systèmes NILM. Ils peuvent identifier les appareils en marche et en arrêt, de ce fait à haute puissance. Cependant, les appareils à faible consommation d'énergie sont plus difficiles à distinguer en utilisant uniquement les paramètres PQ, car ils présentent un chevauchement dans l'espace correspondant [63].

II-2-3 Paramètres PQ combinés avec les paramètres transitoires macroscopiques

Cette combinaison de paramètres est spécifiquement adaptée à l'identification des appareils ayant des pics importants dans leur consommation d'énergie dans un temps transitoire relativement long [17], suivis de variations avec des changements plus lents [26]. Ces variations transitoires sont caractérisées par les bords et les pente sou par leurs profils de puissance [17]. Bien que les caractéristiques transitoires macroscopiques soient relativement simples à obtenir (en utilisant des mesures à basse fréquence), elles ne permettent pas de discriminer les appareils non linéaires ou les appareils multi-états.

II-2-4 Paramètres PQ combinés avec le courant I et la tension V

La combinaison des paramètres, basée sur le courant et la tension électriques (c.-à-d. I_{\max} , V_{\max} , facteur de puissance (FP) et le déphasage) avec les paramètres PQ, montre de bonnes performances pour l'identification des appareils de cuisine utilisant la reconnaissance en temps réel des appareils électriques. Cependant, la détection d'appareil multi-états nécessite plus de paramètres par rapport au cas de l'utilisation des paramètres d'état stationnaire et macroscopiques [64]. Dans la référence [65], le paramètre FP a été également combiné avec les paramètres PQ.

II-2-5 Paramètres PQ combinés avec les paramètres transitoires macroscopiques et les paramètres harmoniques

La prolifération de l'utilisation des appareils électriques non linéaires dans les milieux résidentiels et commerciaux a motivé l'utilisation du contenu harmonique des signaux électriques comme élément discriminant. Sultanem [63] est le premier qui a combiné les harmoniques avec des paramètres basés sur les basses fréquences. Il a également proposé des paramètres PQ et des paramètres transitoires macroscopiques, mais son algorithme est différent du travail de Hart.

II-2-6 Paramètres P-Q combinés avec les harmoniques

Dans la plupart des cas, les harmoniques des formes d'onde du courant électrique $i(t)$ sont combinées avec les paramètres PQ. Dans la référence [38], Srinivasan et al ont utilisé l'amplitude et la phase des 8 premières harmoniques en régime transitoire et en régime permanent. Dans [66], Laughman et al ont utilisé les harmoniques du signal transitoire avec les paramètres PQ en constatant de ce fait que la troisième harmonique peut améliorer la discrimination entre les ordinateurs et les ampoules à incandescence.

II-2-7 Enveloppe spectrale

Dans [66], les auteurs ont proposé un système IAED basé sur la transformée de Fourier à court terme STFT (Short-Time Fourier Transform) appliquée au signal du courant électrique pour calculer des enveloppes spectrales qui résument le contenu harmonique variant dans le temps. Contrairement à la FFT, dans laquelle les informations temporelles sont perdues lors de la transformation du signal dans le

domaine fréquentiel, la transformée STFT utilise une fenêtre fixe pour transformer chaque segment du signal de courte durée en préservant les informations du temps et de fréquence. Cependant, l'utilisation d'une taille de fenêtre fixe pour toutes les fréquences limite la flexibilité de STFT [67].

II-2-8 Paramètres basés sur la transformée en ondelettes

Par rapport à la STFT, la transformée en ondelettes (TO) est une approche plus flexible pour représenter un signal variable. La TO décompose un signal en temps et en échelle à l'aide d'ondelettes avec des propriétés d'échelle adaptables. De cette façon, on peut utiliser des fenêtres plus longues dans le cas où des informations basses fréquences plus précises sont nécessaires. En revanche, des fenêtres plus courtes utilisées dans le cas où des informations hautes fréquences précises sont nécessaires. Chan et al. [28] ont franchi la première étape de la construction d'une signature d'appareil à l'aide de la transformée en ondelettes continues CWT (Continuous Wavelet Transform) du signal de courant électrique. Cette signature est construite à l'aide d'une transformée en ondelettes utilisant l'ondelette mère Daubechies (DB4). Duarte et al. [68] ont ensuite comparé la STFT et la CWT dans la décomposition des parties transitoires de la tension de commutation et ont recommandé CWT comme une approche prometteuse pour extraire les paramètres transitoires dans les systèmes NILM. Pour éviter la complexité de calcul de la CWT, une transformée en ondelettes discrète DWT (Discrete Wavelet Transform) a été utilisée comme alternative pour extraire des paramètres pour les applications NILM. Dans la référence [69], Su et al ont utilisé les coefficients de la DWT en régime transitoire du courant d'activation comme un ensemble de paramètres et ont montré ses avantages par rapport à la STFT dans l'analyse transitoire des appareils. Dans la référence [70], [37], Chang et al ont complété les paramètres PQ avec l'énergie transitoire et le temps de réponse transitoire des formes d'onde de puissance (calculé sur la base des coefficients DWT avec des ondelettes DB3) et ont démontré que la combinaison des paramètres améliore la précision et le temps d'apprentissage du système. Dans [71], Gray et al ont utilisé l'énergie des coefficients d'ondelettes, obtenus au niveau des détails, comme un ensemble de paramètres et ont comparé la précision de classification en utilisant ces paramètres par divers ordres d'ondelettes Daubechies (DB). Ils ont montré que les ondelettes DB d'ordre supérieur (DB5 en particulier)

présentent une précision de classification plus élevée. Tabatabaei et al, dans [72], ont calculé également l'énergie des coefficients d'ondelettes dans chaque niveau de la DWT (obtenus en utilisant des ondelettes de Haar). Enfin, Gillis et al, dans [73], ont proposé une nouvelle ondelette spécifiquement conçue pour l'application NILM. Cette ondelette remplace en effet les ondelettes DB conçues typiquement pour d'autres applications telles que les communications et le traitement d'image.

Bien que les coefficients d'harmoniques et d'ondelettes soient très efficaces pour représenter les formes d'onde non linéaires, ils nécessitent une haute fréquence d'échantillonnage (au moins deux fois la fréquence de l'harmonique la plus élevée destinée à l'extraction) et un traitement complexe du signal.

II-2-9 Trajectoires V-I (tension - courant)

Une forme bidimensionnelle de signature des appareils électriques, créée par la trajectoire tension-courant instantanée (V-I), est suggérée pour caractériser les appareils électriques domestiques. La trajectoire V-I a des significations techniques utiles. Une étude indique que la puissance et les caractéristiques mécaniques de certains appareils électromécaniques jouent un rôle critique dans l'identification des appareils électriques. Lam et al. [74] ont introduit des paramètres basés sur les trajectoires V-I qui ont été utilisés pour la conception d'un système IAED. Hassan et al. [75] ont illustré de nouvelles améliorations dans la catégorisation des appareils électriques en extrayant des paramètres supplémentaires des trajectoires V-I. Du et al. [76] ont tracé d'abord les trajectoires V-I à une grille de cellules et ont introduit pour chacune d'elles une valeur binaire. Finalement, Gao et al. [77] ont utilisé l'image binaire susmentionnée de la trajectoire V-I comme ensemble de paramètres.

II-2-10 Courant inactif

Avec l'utilisation croissante des appareils électriques non linéaires et le courant non sinusoïdal, des théories de puissance plus avancées sont envisagées pour la décomposition de la puissance apparente. La décomposition du courant en courants actif ($i_a(t)$) et non actif ($i_f(t)$) ont été proposés par Fryze [78]. Huang et al, dans [79], ont montré que la similarité entre les courants inactifs de divers appareils est plus faible par rapport à leurs formes d'ondes de courant et de puissances instantanées. Cependant,

l'efficacité des paramètres basés sur le courant inactif, dans l'amélioration des performances des algorithmes NILM, n'a pas été testée à l'aide d'un algorithme de discrimination des appareils électriques. Teshome et al. [80] proposent des paramètres basés sur la tension non active $V_f(t)$.

II-2-11 Paramètres non conventionnelles

Patel et al. [2] ont utilisé la FFT du bruit de la tension électrique pour définir les signatures d'appareils et ont étendu leurs travaux pour considérer les coefficients de Fourier des signaux d'interférence électromagnétique dans la gamme 36-500 kHz [81]. Bien que ces paramètres atteignent une précision d'identification élevée, ils sont sensibles au câblage électrique domestique. De plus, ils nécessitent que les appareils soient équipés d'une alimentation à découpage, ce qui n'est pas le cas pour tous les appareils électroménagers. Kim et al. [82] ont combiné des paramètres non électriques tels que la distribution de la durée ON / OFF, la fréquence d'utilisation des appareils et la corrélation entre l'utilisation de divers appareils avec la fonction de puissance réelle pour améliorer la discrimination des appareils électriques. Wang et al. [83] ont représenté les paramètres basés sur la forme de puissance réelle en utilisant deux formes unitaires qui sont les rectangles et les triangles. Dans le même contexte, Koutitas et al. [84] ont remplacé la série temporelle d'un signal de puissance continue par un ensemble d'impulsions discrètes qui sont calculées selon la première dérivée du signal de puissance. Ils ajoutent en outre des informations sur le comportement humain (c'est-à-dire la probabilité d'utilisation, la probabilité des impulsions de voisinage, la durée d'une impulsion, la séquence de fonctionnement et les conditions externes) à ces impulsions.

II-3 Analyse harmonique

Plusieurs travaux de recherches ont utilisé les coefficients harmoniques de la décomposition en série de Fourier comme paramètres discriminants pour la tâche IAED [1]. Cette décomposition est appliquée sur des signaux discrets du courant électrique en tenant compte de leur aspect de quasi-périodicité. La section suivante décrira théoriquement cette analyse.

II-3-1 Décomposition en Série de Fourier (DSF)

La décomposition en série de Fourier (DSF) d'un signal analogique périodique $x(t)$ de période T_0 est donnée, sous sa forme complexe, par:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{j2\pi k f_0 t} \quad \text{avec } f_0 = \frac{1}{T_0} \text{ est la fréquence fondamentale} \quad (II - 1)$$

où C_k , sont les coefficients de Fourier, donnés comme suit:

$$C_k = \frac{1}{T_0} \int_0^{T_0} x(t) e^{-j2\pi k f_0 t} dt \quad (II - 2)$$

Les coefficients C_{-1} et C_1 correspondent à la fréquence fondamentale, alors que Les coefficients C_k (avec $k > 1$) sont appelés les coefficients des harmoniques ($f_k = k f_0$). C_0 est la moyenne du signal.

Pratiquement, la DSF s'applique sur des signaux discrets périodiques échantillonnés avec une fréquence d'échantillonnage f_e .

La DSF d'un signal discret périodique $x(n)$ de période N est donnée comme suit:

$$x[n] = \sum_{k=0}^{N-1} C_k \exp\left(\frac{j2\pi kn}{N}\right) \quad (I - 3)$$

où N est la période de $x[n]$ et C_k sont les coefficients de la série de Fourier exprimés par:

$$C_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \exp\left(\frac{-j2\pi kn}{N}\right) \quad k = 0, 1, \dots, N - 1 \quad (II - 4)$$

Les coefficients C_k peuvent être obtenus en utilisant la Transformée de Fourier Discrète DFT (Discrete Fourier Transform):

$$C_k = \frac{X(k)}{N}, \quad k = 0, 1, \dots, N - 1 \quad (II - 5)$$

Où la DFT $X(k)$ est donnée par:

$$X(k) = \sum_{n=0}^{N-1} x[n] \exp\left(\frac{-j2\pi kn}{N}\right) \quad k = 0, 1, \dots, N-1 \quad (II-6)$$

Dans le cas des signaux réels, le spectre d'amplitude est pair. Ainsi, seulement les $N/2$ premiers coefficients sont pris en considération.

Dans le domaine NILM, la DSF est effectuée sur des fenêtres temporelles chevauchées de courtes durées dans lesquelles le signal est considéré stationnaire. Le processus de calcul des coefficients C_k est couramment connu sous le nom STFS (Short-Time Fourier Series). Généralement, le module des coefficients est pris en considération dans l'étape d'extraction des paramètres pour la tâche IAED.

II-3-2 Application de l'analyse harmonique dans les systèmes IAED

L'analyse harmonique est l'une des techniques les plus populaires dans le domaine du traitement du signal. Cette technique est utilisée également dans la conception des systèmes IAED, plus particulièrement dans l'étape d'extraction des paramètres. Dans [1], les auteurs ont utilisé la méthode d'extraction des paramètres STFS, appliquée au signal du courant électrique, pour la conception d'un système d'identification des appareils électriques basée sur le classificateur HMM. Ce système est illustré sur la figure II-1.

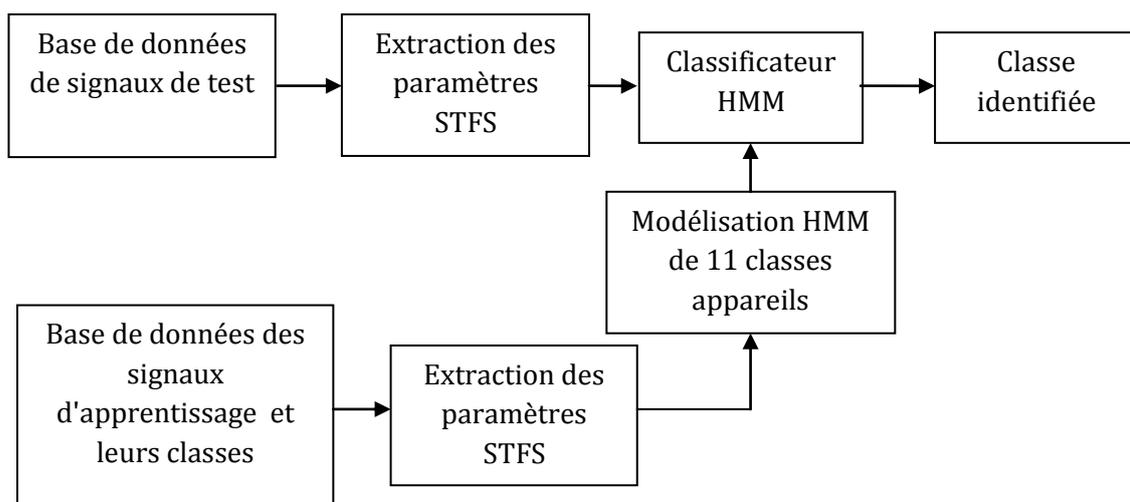


Figure II-1. Schéma du système d'identification proposé dans [1]

Dans ce travail, les auteurs ont proposé de concevoir un système d'identification d'appareils électriques basé sur la modélisation des différentes classes d'appareils électriques par des modèles de Markov cachés HMM (Hidden Markov Model) associés aux modèles GMM (Gaussian Mixture Model). Plusieurs expériences ont été réalisées pour chercher la bonne configuration, en termes de précision, du système.

Les performances du système sont évaluées sous la base de données PLAID qui se caractérise par un grand nombre d'installation électrique (55 maisons) permettant l'enregistrement des signaux de différents appareils. La répartition de cette base est effectuée en prenant des exemples de signaux d'appareils de chaque maison dans la base d'apprentissage et la base de test.

Chaque signal du courant électrique est segmenté, en se basant sur la variation de l'énergie, en une partie transitoire et une partie permanente. Cette segmentation a pour objectif d'étudier l'influence de ces deux parties sur les performances du système.

Les résultats les plus importants de cette conception peuvent être résumés par les points suivants:

- La partie transitoire du signal est plus informative que la partie permanente;
- Les coefficients harmoniques correspondant à l'ordre impair sont plus pertinents que ceux correspondant à l'ordre pair;
- La prise d'un sous-ensemble réduit de paramètres STFS permet d'obtenir de meilleures performances en comparaison avec le cas de l'utilisation de tous les paramètres STFS.

II-3-3 Approche d'extraction proposée

Notre approche proposée consiste à représenter une sous-séquence de N vecteurs, composé chacun de M paramètres STFS, par un seul vecteur de $4 M$ paramètres statistiques en estimant la moyenne, l'écart type, l'asymétrie et l'aplatissement de chaque paramètre harmonique. Pratiquement, chaque sous-séquence de paramètres harmoniques est représentée par une matrice appelée $MH(M \times N)$ dont chaque ligne représente un paramètre harmonique et chaque colonne représente un vecteur instantané ou une fenêtre d'analyse. Chaque paramètre statistique de paramètre

harmonique d'ordre $i(i=1, 2, \dots, M)$ est estimé à partir des N valeurs de la ligne i de la matrice MH [5]. Ces paramètres statistiques sont donnés comme suit:

- **La moyenne μ du paramètre harmonique p d'ordre i :**

$$\mu(p_i) = E[MH(i, 1:N)] \quad (IV - 7)$$

$$\mu(p_i) = \frac{1}{N} \sum_{j=1}^N MH(i, j) \quad i = 1, \dots, M \quad (IV - 8)$$

- **L'écart type σ :**

$$\sigma(p_i) = \sqrt{E [(MH(i, 1:N) - \mu(p_i))^2]} \quad (IV - 9)$$

estimé par :

$$\sigma(p_i) = \sqrt{\frac{1}{N} \sum_{j=1}^N (MH(i, j) - \mu(p_i))^2} \quad i = 1, \dots, M \quad (IV - 10)$$

- **L'asymétrie (skewness Sk) :**

$$Sk(p_i) = E \left[\left(\frac{MH(i, 1:N) - \mu(p_i)}{\sigma(p_i)} \right)^3 \right] \quad (IV - 11)$$

estimée par :

$$Sk(p_i) = \frac{\frac{1}{N} \left(\sum_{j=1}^N (MH(i, j) - \mu(p_i))^3 \right)}{(\sigma(p_i))^3} \quad i = 1, \dots, M \quad (IV - 12)$$

- **L'aplatissement (kurtosis Ku) :**

$$Ku(p_i) = E \left[\left(\frac{MH(i, 1:N) - \mu(p_i)}{\sigma(p_i)} \right)^4 \right] \quad (IV - 13)$$

estimé par :

$$Ku(p_i) = \frac{\frac{1}{N} \left(\sum_{j=1}^N (MH(i, j) - \mu(p_i))^4 \right)}{(\sigma(p_i))^4} \quad i = 1, \dots, M \quad (IV - 14)$$

Enfin, chaque courant électrique d'un appareil est représenté par une séquence VS de K vecteurs statistiques $VS_k (k = 1:K)$ tel que :

$VS = [Vs_1, Vs_2, \dots, Vs_K]$, où K est le nombre de vecteurs statistiques.

Vs_k est la concaténation des paramètres statistiques des M paramètres harmoniques comme suit:

$$Vs_k = [\mu(p_1) \ \sigma(p_1) \ Sk(p_1) \ Ku(p_1) \dots \mu(p_M) \ \sigma(p_M) \ Sk(p_M) \ Ku(p_M)] \quad (IV - 15)$$

Ainsi, chaque signal, représenté initialement par $N \times M \times K$ paramètres, est converti en $M \times K$ paramètres. Cette conversion permet d'améliorer la complexité en termes de calcul et d'espace mémoire en réduisant la taille de la représentation initiale de N fois [5].

II-4. Transformée de Fourier à court terme STFT

La DSF peut être généralisée aux signaux non périodiques en utilisant la notion de la Transformée de Fourier (TF). La TF d'un signal analogique $x(t)$ est définie comme suit :

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f t} dt \quad (II - 16)$$

Où $X(f)$ représente le spectre du signal $x(t)$.

$x(t)$ est donnée, en utilisant la TF inverse, comme suit :

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi f t} df \quad (II - 17)$$

Pratiquement, le spectre du signal est déterminé en utilisant la DFT. Rappelons que la DFT est donnée par la formule (II-5) dans laquelle k variant de 0 à $N-1$, représente les composantes fréquentielles $\frac{k}{N}$. La DFT du signal $x(t)$ est calculée généralement en utilisant l'algorithme FFT (Fast Fourier Transform).

Dans le cas d'un signal non stationnaire, il est indispensable de le représenter dans l'espace-temps-fréquence en utilisant la STFT. La STFT d'un signal discret $x(n)$ est définie comme suit :

$$g(f, m) = \sum_{n=-\infty}^{+\infty} x[n] w(n-m) e^{-2j\pi f n} \quad (II-18)$$

Où $w(n-m)$ est la fonction fenêtre glissante de m points.

Plusieurs choix de la fonction w peuvent être envisagés tels que la fonction rectangulaire, la fonction triangulaire, la fonction Gauss et la fonction Hamming. La largeur de cette fenêtre doit être choisie en assurant une quasi-stationnarité du signal.

Pratiquement, la STFT est calculée en discrétisant la fréquence f en N points fréquentiels :

$$g(m, k) = \sum_{n=-\infty}^{+\infty} x[n] w(n-m) e^{-\frac{2j\pi kn}{N}} \quad k = 0, \dots, N-1 \quad (II-19)$$

Le choix de la taille de la fenêtre affecte la résolution dans le domaine temporel et le domaine fréquentiel [85].

L'inconvénient majeur de la STFT est que la largeur de la fenêtre glissante est fixe. L'outil idéal serait une fenêtre qui s'adapte aux variations de la fréquence dans le signal à analyser. Cet outil, qui existe depuis plusieurs dizaines d'années, s'appelle l'analyse en ondelettes. Dans la section suivante, nous présentons cette variante d'analyse.

II-5 Analyse par Ondelettes

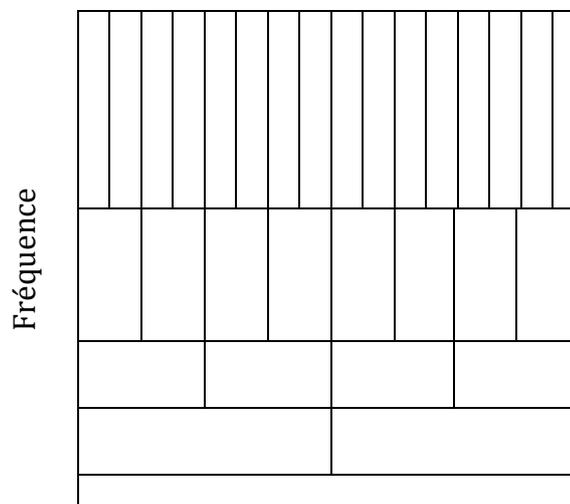
La STFT utilise une fenêtre glissante pour trouver le spectrogramme donnant des informations de temps et de fréquence sur le signal. Cependant, la largeur de la fenêtre limite la résolution en fréquence. Une méthode alternative, avec des propriétés intéressantes, peut être utilisée pour donner de meilleurs résultats. Il s'agit de l'analyse en ondelettes évoquée pour la première fois par Alfred Haar en 1909. C'est une représentation des signaux qui permet la mise en valeur simultanée des informations temporelles et fréquentielles. Un avantage de l'analyse en ondelettes est la capacité d'effectuer une analyse locale basée sur de petites ondelettes de durées limitées [86].

II-5-1 Historique des ondelettes

Alfred Haar a été le premier à mentionner les ondelettes en 1909 dans sa thèse de doctorat. Cependant, le concept de l'ondelette n'existait pas à l'époque. En effet, ce n'est qu'à partir de l'année 1981 que Paul Levy a trouvé la fonction de base de Haar à échelle variable supérieure aux fonctions de base de Fourier. La méthode de transformation qui consiste à décomposer un signal en coefficients d'ondelettes et à reconstruire à nouveau le signal d'origine a été dérivée en 1981 par Jean Morlet et Alex Grossman. Par la suite, Morlet et le physicien Alex Grossman ont inventé le terme ondelettes en 1984. En 1985 le mathématicien Yves Meyer a construit la deuxième ondelette orthogonale appelée ondelette de Meyer. En 1986, Stéphane Mallat et Yves Meyer ont développé une analyse multi-résolutions à l'aide des ondelettes. Ils ont mentionné la fonction d'échelle des ondelettes pour la première fois ce qui permis de plus en plus aux chercheurs et aux mathématiciens de construire leurs propres familles d'ondelettes. Vers l'année 1998, Ingrid Daubechies a trouvé une méthode systématique pour construire l'ondelette orthogonale (de support compact) par l'utilisation de la théorie de l'analyse par ondelettes multi-résolution dans le but de construire sa propre famille d'ondelettes. En 1989, Mallat a proposé la transformation rapide en ondelettes. Avec l'apparition de cet algorithme rapide, la transformée en ondelettes a eu de nombreuses applications dans le domaine du traitement du signal.

II-5-2 Transformée en ondelette continue (TOC)

En utilisant une approche appelée analyse multi-résolution, il est possible d'analyser un signal à différentes fréquences avec différentes résolutions. Le changement de résolution est schématisé sur la Figure II-2.



 Temps

Figure II-2 Plan multi-résolution temps-fréquence

On remarque, d'après cette figure, que la largeur de la fenêtre pour l'analyse en ondelette est modifiée au fur et à mesure qu'on calcule individuellement la transformée pour chacune des composantes spectrales. Ce dernier point représente la caractéristique la plus significative de la transformée en ondelettes. En effet, la transformée en ondelettes a la capacité de changer la taille de la fenêtre à différentes résolutions de temps-fréquence comme représenté sur la Figure II-2. Les fonctions de base d'ondelettes peuvent être dilatées ou comprimées dans la fenêtre d'analyse. Pour la résolution de la Figure II.2, il est supposé que les basses fréquences durent pendant toute la durée du signal, tandis que les hautes fréquences apparaissent de temps en temps avec une courte durée. C'est souvent le cas dans les applications pratiques. La fonction d'analyse des ondelettes $\psi(t)$ est également appelée ondelette mère.

La transformée en ondelettes continue TOC d'un signal analogique $x(t)$ est exprimée en fonction de l'échelle s et la translation (position) τ , par :

$$C(s, \tau) = \langle x(t), \psi_{s,\tau}(t) \rangle$$

$$= \int_{-\infty}^{+\infty} x(t) \psi_{s,\tau}^*(t) dt \quad (II - 20)$$

où $\psi(t)$ est la fonction d'ondelette mère. La fonction $\psi_{s,\tau}(t)$ est la fonction fille obtenue par translation et dilatation de la fonction mère $\psi(t)$.

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (II - 21)$$

Avec :

τ : Facteur de translation ;

s : Facteur de dilatation « Echelle » ;

$\frac{1}{\sqrt{s}}$: Facteur de normalisation de l'énergie.

Le signal $x(t)$ peut être construit, en utilisant la transformée TOC, comme suit:

$$x(t) = \frac{1}{K_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{|s|^2} c(s, \tau) \psi_{s,\tau}(t) ds d\tau \quad (II - 22)$$

$$\text{Où :} \quad K_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega \quad (II - 23)$$

$\hat{\Psi}(\omega)$ est la transformée de Fourier de la fonction $\psi(t)$.

$\psi(t)$ est classée comme une ondelette analysante si les critères mathématiques suivants sont satisfaits [87]:

- L'ondelette $\psi(t)$ doit avoir une moyenne nulle (équation II-24) et une énergie finie (équation II-25). Ces deux propriétés sont données comme suit :

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (II - 24)$$

$$E = \int_{-\infty}^{+\infty} |\psi(t)|^2 dt < \infty \quad (II - 25)$$

- Si $\Psi(f)$ est la transformée de Fourier de l'ondelette $\psi(t)$, la condition d'admissibilité suivante doit être vérifiée :

$$C_\psi = \int_0^{+\infty} \frac{|\Psi(f)|^2}{f} df < \infty \quad (II - 26)$$

Où : $\Psi(0)=0$ (pas de composante de fréquence nulle) et la valeur de C_ψ dépendent de l'ondelette choisie.

La figure II-3 montre un exemple de différentes échelles d'une ondelette connue sous le nom d'ondelette de Morlet [88].

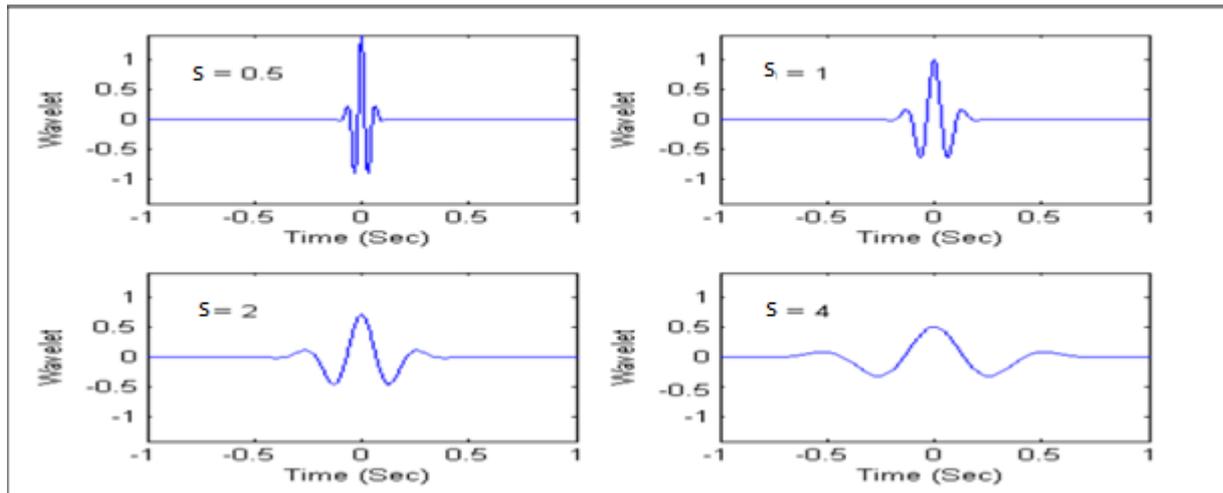


Figure II-3 Ondelette de Morlet avec différentes échelles ($s=0.5$, $s=1$, $s=2$, $s=4$) [89]

La figure II-4 illustre différentes translations de la fonction d'ondelette Morlet pour une échelle fixe.

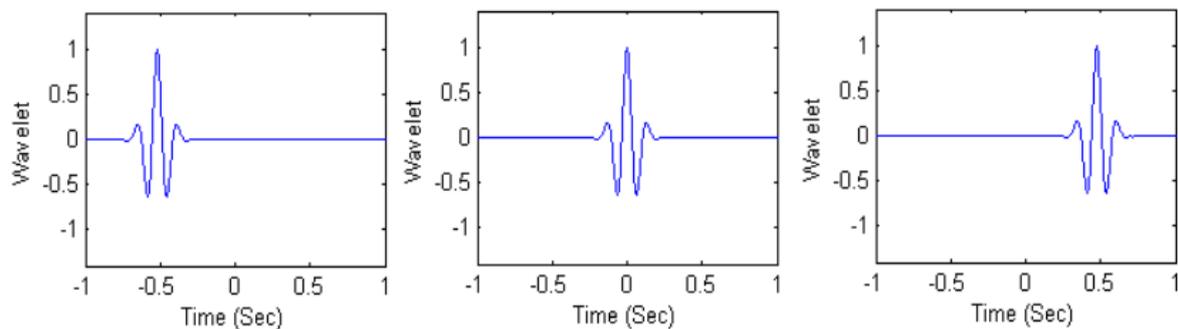


Figure II-4 Ondelette de Morlet avec différents facteurs de translation ($\tau = -0.5$, $\tau = 0$ et $\tau = 0.5$) [89]

Dans la littérature scientifique, plusieurs familles spéciales de fonctions d'ondelettes sont développées [86], [90], [91]. Ces ondelettes sont compactes, orthogonales et se caractérisent par des filtres d'analyse passe-bas et passe-haut. Les familles d'ondelettes mères les plus connues sont les ondelettes Daubechies (Haar ou db1,db2,...), les ondelettes Symlet (sym1,sym2,...) et les ondelettes Coiflet (coif1,coif2,...). Les différentes familles d'ondelettes sont utilisées selon leurs propriétés et leurs applications. La Figure (II-5) illustre quelques fonctions d'ondelettes mères des familles Daubechies, Coiflet et Symlet.

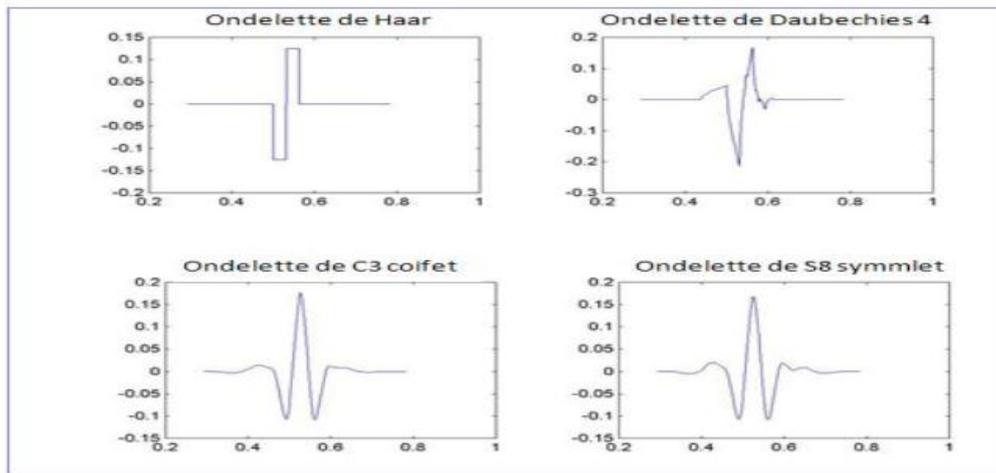


Figure II-5 Quelques familles d'ondelettes [89]

Le paramètre échelle est utilisé en analyse par ondelettes, de façon que les basses fréquences (grandes échelles) fournissent une information globale sur le signal (habituellement sur tout l'étendu du signal) alors que les hautes fréquences (faibles échelles) donnent des informations détaillées sur un motif caché dans le signal (généralement de faible durée). La notion de fréquence est remplacée donc par la notion d'échelle en analysant les informations de basses fréquences et celles de hautes fréquences, respectivement par la dilatation et contraction de l'ondelette de mère.

II-5-3 Transformée en ondelette discrète

La transformée d'ondelette discrète (TOD) décompose le signal en un ensemble d'ondelettes mutuellement orthogonales. Elle consiste à discrétiser le facteur d'échelle s ainsi le facteur de translation τ . Ces coefficients s et τ seront discrétisés de la manière suivante:

$$s = s_0^j \text{ et } \tau = k\tau_0 s_0^j \quad \text{avec } s_0 > 1 \text{ et } \tau_0 > 0 \text{ sont fixés et appartiennent à } \mathbb{Z}$$

La formule de la fonction mère devient alors :

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \psi\left(\frac{t - ks_0^j \tau_0}{s_0^j}\right) \quad (\text{II} - 27)$$

La formule de la TOD d'un signal $x(t)$ est donnée par :

$$C_a(j, k) = \frac{1}{s_0^j} \int_{-\infty}^{+\infty} x(t) \psi \left(\frac{t - k s_0^j \tau_0}{s_0^j} \right) dt \quad (II - 28)$$

Généralement, une échelle dyadique est donnée par: $s_0 = 2$ et $\tau_0 = 1$. Ainsi, les équations (II-27) et (II-28) deviennent:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - k 2^j}{2^j} \right) \quad (II - 29)$$

$$C_a(j, k) = \frac{1}{2^j} \int_{-\infty}^{+\infty} x(t) \psi \left(\frac{t - k 2^j}{2^j} \right) dt \quad (II - 30)$$

La TOD a été élaborée dans le cadre de l'analyse multi-résolution basée sur deux types de filtres numériques passe-bas h et passe-haut g .

Les coefficients d'ondelettes discrètes peuvent être obtenus à chaque niveau j , par la formule suivante :

$$\begin{cases} a_j[n] = \sum_l a_{j-1}[l - 2n] h[l] \\ d_j[n] = \sum_l a_{j-1}[l - 2n] g[l] \end{cases} \quad (II - 31)$$

Notez que l'ondelette mère n'apparaît pas directement dans ces expressions récursives mais sa forme d'onde continue peut-être récupérée à partir de la séquence g [1]. De même, une autre forme d'onde continue (dite fonction de mise à l'échelle) peut être extraite de la séquence h [1].

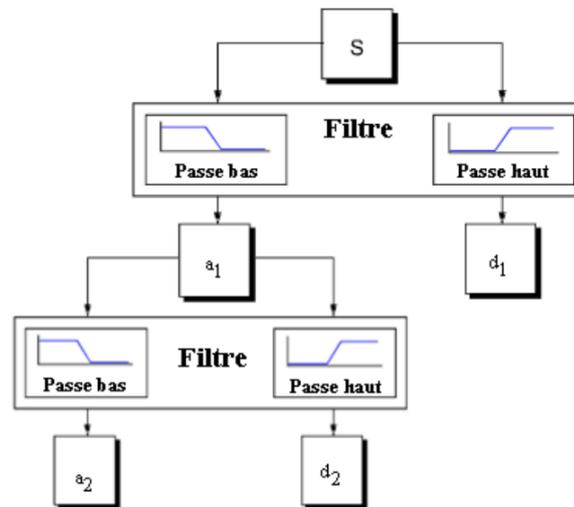


Figure II-6– Schéma de la décomposition en ondelettes [92]

L'algorithme de la décomposition en ondelettes est initialisé au niveau 0 en fixant $a_0[n] = x[n]$ défini sur N échantillons. A chaque itération, les filtres divisent la bande passante de données en bandes basse fréquence et haute fréquence (le résultat peut donc être sous-échantillonné d'un facteur égal à 2 qui est le facteur d'échelle dyadique). Les composantes basses fréquences sont ainsi représentées par les coefficients d'approximation $a_j[n]$ tandis que les composantes hautes fréquences sont représentées par les coefficients de détail $d_j[n]$. Les coefficients d'ondelettes discrètes TOD, au niveau de décomposition p ; peuvent être placés dans un vecteur comme la concaténation des coefficients de détail (calculés à toutes les échelles) plus les coefficients d'approximation restants (calculés à l'échelle p) $\{d_1[n], d_2[n] \dots d_p[n], a_p[n]\}$. En raison du sous-échantillonnage de facteur 2, le nombre de coefficients $d_j[n]$ à l'itération j est $N_j = \frac{N}{2^j}$. Cela signifie que les N échantillons de $x[n]$ est conservés dans le domaine TOD avec N coefficients. La profondeur de la décomposition maximale peut être $\log_2(N)$ mais elle dépend pratiquement de la longueur des filtres [71].

II-5-4 Application de l'analyse par ondelettes dans les systèmes IAED

Dans [40], les auteurs ont proposé un système IAED utilisant les modèles HMM et l'extraction des paramètres basée sur l'analyse par ondelettes. Cette analyse permet d'obtenir une représentation plus compacte par comparaison avec l'analyse SFTS. En

effet, trois descripteurs de paramètres à savoir le DWE (Discrete Wavelet Energy), le LWE (Log Wavelet Energy) et le WCC (Wavelet Cepstral Coefficients) ont été extraits à partir de cette analyse. La pertinence des paramètres de ces descripteurs a été étudiée à l'aide d'une procédure de sélection de paramètres de type "Wrappers". La figure II-7 montre le schéma synoptique pour l'extraction des différents descripteurs utilisés dans la référence [40].

Courant

électrique

Figure II-7: Processus d'extraction des paramètres DWE/LWE/WCC avec fenêtrage de Hamming [40]

Dans la référence [40], le premier descripteur de paramètres DWE consiste à calculer l'énergie normalisée par rapport à l'énergie totale E_T , à partir des coefficients de détails de chaque niveau de décomposition. Ainsi, le vecteur de paramètres extraits est constitué de l'ensemble des énergies normalisées des différents niveau de décomposition et de l'énergie normalisée calculée à partir des coefficients d'approximation du dernier niveau de décomposition p .

L'ensemble des N coefficients d'ondelettes (N étant le nombre d'échantillons) est réduit à un ensemble de $p + 1$ paramètres donnés comme suit :

$$\begin{cases} E(d_j) = \frac{1}{E_T} \sum_{n=0}^{n=N_j-1} |d_j[n]|^2 \\ E(a_p) = \frac{1}{E_T} \sum_{n=0}^{n=N_p-1} |a_p[n]|^2 \end{cases} \quad \text{pour } j = 1 \dots p, \quad N_j = \frac{N}{2^j} \quad (\text{II} - 32)$$

Où :

N_j : est le nombre de coefficients d'ondelettes qui sont calculés à chaque niveau j ;

$a_p [n]$: sont les coefficients d'approximation ;

$d_j [n]$: sont les coefficients de détail au niveau j .

Les valeurs d'énergie sont normalisées par rapport à l'énergie totale E_T . Cette méthode de calcul des énergies normalisées des coefficients d'ondelettes DWE (Discrete Wavelet Energy) nous permet d'avoir une bonne réduction de dimensionnalité [71].

Le calcul des coefficients LWE (Log Wavelet Energy) consiste à appliquer le logarithme sur chacun des paramètres DWE. Ces coefficients sont donnés par :

$$\begin{cases} LE(d_j) = \log \left(\frac{1}{E_T} \sum_{n=0}^{n=N_j-1} |d_j [n]|^2 \right) \\ LE(a_p) = \log \left(\frac{1}{E_T} \sum_{n=0}^{n=N_p-1} |a_p [n]|^2 \right) \end{cases} \text{ pour } j = 1 \dots p, N_j = \frac{N}{2^j} \quad (\text{II} - 33)$$

Le calcul des coefficients WCC (Wavelet Cepstral Coefficients) consiste à appliquer, sur chaque paramètre LWE, la Transformée en Cosinus Discrète (TCD) en prenant la valeur réelle (RDCT). Ces coefficients sont donnés par :

$$\begin{cases} WCC(d_j) = RDCT \left[\log \left(\frac{1}{E_T} \sum_{n=0}^{n=N_j-1} |d_j [n]|^2 \right) \right] \\ WCC(a_p) = RDCT \left[\log \left(\frac{1}{E_T} \sum_{n=0}^{n=N_p-1} |a_p [n]|^2 \right) \right] \end{cases} \text{ pour } j = \dots .. p \quad (\text{II} - 34)$$

Les résultats obtenus de ce processus d'extraction des paramètres montrent que le descripteur WCC extrait sur des fenêtres d'analyse de durée 8 périodes de durée 1/60 s présente des performances supérieures par rapport aux descripteurs STFS, DWE et LWE. Des recommandations sont également données pour sélectionner les paramètres suivants :

- La famille d'ondelettes;
- L'ordre des ondelettes mères;
- Le niveau de décomposition;
- Les paramètres pertinents.

Les meilleurs résultats sont obtenus en choisissant l'ondelette Daubechies 4 et le niveau de décomposition 6 (ou l'ondelette Coiflet 2 et le niveau de décomposition 7).

Dans notre travail, on propose de calculer les paramètres statistiques des paramètres de différents descripteurs (DWE, LDW et WCC) en utilisant la même stratégie de calcul des paramètres statistiques des coefficients harmoniques (Voir section II-3-3) [6]. On propose également de calculer le logarithme d'énergie sur chaque fenêtre d'analyse pour l'estimation d'un nouveau paramètre statistique que nous nommons LOG_E [7]. Le paramètre du Logarithme d'Energie (LE) est largement utilisé avec les coefficients MFCC (Mel-Frequency Cepstral Coefficients) pour la tâche de la reconnaissance automatique de la parole (RAP) [93]. Il est considéré comme l'un des paramètres les plus pertinents pour la tâche de RAP [45]. Ce paramètre consiste à appliquer le logarithme sur l'énergie totale E calculée sur le segment du signal de la fenêtre d'analyse de N échantillons, comme suit:

$$E = \sum_{n=0}^{N-1} x(n)^2 \quad (\text{II} - 35)$$

Dans notre travail, le paramètre statistique LOG_E est estimé en utilisant la même stratégie de calcul des paramètres statistiques décrite précédemment.

La concaténation de ce descripteur avec les descripteurs basés sur l'analyse par ondelettes sera étudiée dans le chapitre (IV).

II-6 Conclusion

Dans ce chapitre, nous avons donné un état de l'art approfondi sur les techniques d'extraction des paramètres utilisées dans les systèmes IAED. En effet, deux approches d'analyse ont été présentées. Nous avons montré que la première approche, basée sur l'analyse harmonique souffre de la grande dimensionnalité des vecteurs de paramètres et de la taille de la séquence des vecteurs. Nous avons montré aussi qu'une solution, proposée pour minimiser cette taille, consiste à extraire des paramètres statistiques représentant une sous-séquence de vecteurs. De plus, une autre solution, permettant de réduire la dimension des vecteurs, consiste à sélectionner les paramètres les plus pertinents.

Concernant la deuxième approche d'analyse qui consiste à appliquer la transformée d'Ondelette Discrète pour extraire des paramètres énergétiques, nous avons montré qu'elle permet d'avoir une représentation plus compacte que celle basée sur l'analyse harmonique. En outre, dans le but de réduire la taille de la séquence des vecteurs, la stratégie de calcul des paramètres statistiques doit être appliquée.

Des applications de ces approche d'analyse, dans les systèmes IAED, ont été décrites en présentant les principes des algorithmes proposés dans des travaux liés à ces approches.

CHAPITRE III:
SELECTION DES PARAMETRES
PERTINENTS POUR LA
CLASSIFICATION

III-1 Introduction

Les systèmes IAED exigent une étape d'extraction des paramètres permettant de convertir chaque signal en une séquence de vecteurs de paramètres. Plusieurs méthodes d'extraction sont décrites dans le chapitre II. La méthode d'extraction, basée sur l'analyse harmonique, souffre de la grande dimension des vecteurs de paramètres. Une des solutions pour remédier à ce problème est de chercher une autre méthode d'extraction telle que l'analyse par ondelette qui permet d'extraire des vecteurs de paramètres avec une représentation plus compacte (faible dimension). Une autre solution, utilisant l'analyse harmonique, consiste à réduire la dimensionnalité des vecteurs de paramètres en appliquant des algorithmes de sélection. Dans notre travail en s'intéresse à la sélection de paramètres.

La sélection des paramètres (caractéristiques, attributs, variables) pertinents est une étape importante dans les systèmes de reconnaissance de formes. En effet, elle permet de réduire la dimensionnalité des vecteurs de paramètres. Elle consiste à sélectionner le plus petit sous ensemble de paramètres pertinents à partir d'un ensemble initial de paramètres sans transformation (en préservant l'interprétation).

Les objectifs supplémentaires de la sélection de paramètres sont donnés comme suit:

- L'amélioration de la complexité en termes de temps de calcul et d'espace mémoire ;
- L'amélioration probable de la précision du système en évitant le phénomène de la malédiction de la dimensionnalité ;
- L'atteinte d'une compréhension meilleure et plus simple du processus qui génère les données (les paramètres) [94], [95].

Cette tâche de sélection est difficile en raison du problème d'interaction entre les paramètres dans le cas de grande dimension et du grand nombre de sous-ensembles. Subséquemment, plusieurs méthodes heuristiques ont été proposées pour remédier à ces problèmes.

En fonction de l'objectif visé et des ressources disponibles, le processus de sélection peut être réalisé de plusieurs façons. La réduction de l'ensemble des paramètres est basée sur les mesures de pertinence et de redondance de ces derniers pour la tâche de

classification ou celle de la régression. Le but de la sélection des paramètres est de maximiser la pertinence et de minimiser la redondance. Plus précisément, un paramètre est généralement classé comme : 1) fortement pertinent, 2) faiblement pertinent, non redondant, 3) non pertinent et 4) redondant [96], [97]. Un paramètre fortement pertinent est toujours nécessaire pour un sous-ensemble de paramètres pertinents [96]. Dans certaines conditions, un paramètre faiblement pertinent n'est pas toujours nécessaire pour un sous-ensemble optimal.

Dans de nombreuses applications, la très grande taille de l'ensemble de données risque de ne pas bien fonctionner avant de supprimer les paramètres non pertinents. En fait, il est montré pratiquement que la détérioration des performances des classificateurs peut être causée par l'augmentation du nombre de paramètres et l'utilisation d'une base de données d'entraînement limitée [45]. Ce phénomène, appelé la malédiction de la dimensionnalité, provoque le phénomène de peaking [45]. Tandis que, la réduction du nombre de paramètres non pertinents/redondants réduit considérablement le temps d'exécution des algorithmes d'apprentissage et des algorithmes de classification [94], [98].

III-2 Définitions de la sélection des paramètres

Dans la littérature scientifique, la sélection des paramètres est définie par de nombreux auteurs. Cependant, beaucoup d'entre eux possèdent des définitions similaires. Celles qui sont conceptuellement différentes sont données comme suit:

- Kira, K. et al dans [99]: La sélection des paramètres consiste à trouver le sous-ensemble de paramètres de taille minimale nécessaire et suffisant pour un système basé sur l'apprentissage automatique.
- Narendra, P.M et al dans [100]: La sélection des paramètres consiste à sélectionner un sous-ensemble de M paramètres parmi un ensemble initial de N paramètres, tel que $M < N$, de sorte que la valeur d'une fonction critère soit optimisée sur tous les sous-ensembles de taille M .
- Koller, D. et al dans [101]: La sélection des paramètres consiste à choisir un sous-ensemble de paramètres sélectionnés pour diminuer la taille de l'ensemble

initial de paramètres sans diminuer de manière significative la précision de classification.

- Kohavi, R. et al dans [102] : La sélection des paramètres consiste à sélectionner un petit sous-ensemble de telle sorte que la distribution de classes résultante en utilisant les paramètres sélectionnés, soit aussi proche que possible de la distribution de classes en utilisant tous les paramètres .
- M. Dash et al dans [103]: Dans cette référence, la définition de la sélection des paramètres tient compte des deux facteurs cités dans [101] et [102]. En effet, la sélection des paramètres consiste à sélectionner le sous-ensemble de paramètres de taille minimale en fonction des critères suivants : (i) Le taux de classification ne diminue pas de manière significative. (ii) la distribution de classes résultantes, utilisant les paramètres sélectionnés, ressemble à la distribution des classes utilisant l'ensemble initial de paramètres.

Idéalement, les méthodes de sélection de paramètres cherchent à trouver le meilleur sous-ensemble parmi les 2^N possibles. Cependant cette procédure est exhaustive est trop coûteuse et pratiquement exagérée, même pour un ensemble de paramètres de taille moyenne. D'autres méthodes sont basées sur des algorithmes de recherche heuristiques ou aléatoires. Ces méthodes nécessitent un critère d'arrêt pour empêcher la recherche exhaustive des sous-ensembles. Il existe quatre étapes de base dans une méthode de sélection de paramètres typique (voir la Figure III-1) [103].

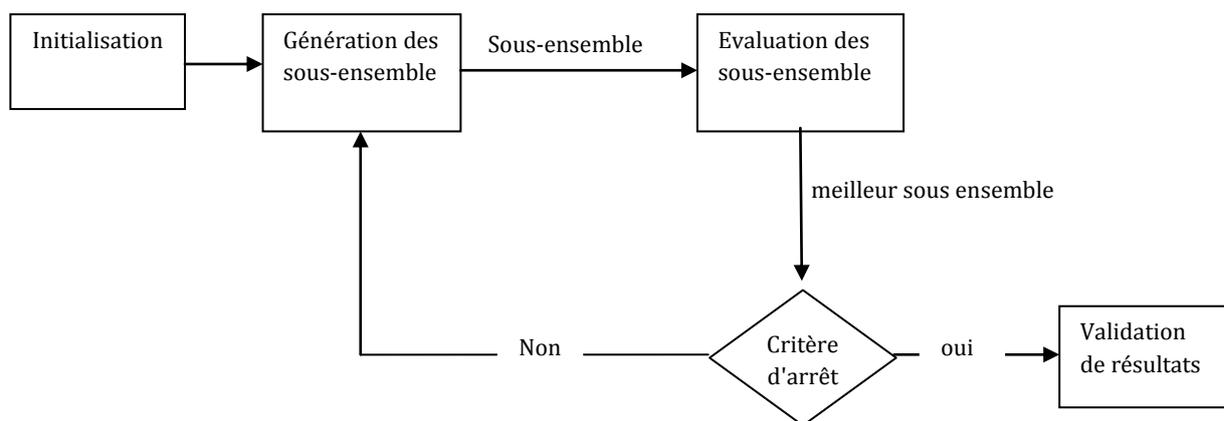


Figure III-1 Processus général de sélection des paramètres [103]

III-3 Classification des méthodes de Sélection

Les méthodes de sélection sont généralement classées en trois groupes principaux: les méthodes "Wrappers", les méthodes "Filters" et les méthodes intégrées (hybrides) [94]. Dans Les méthodes "Wrappers" [95], les performances sont généralement mesurées en termes de taux de classification obtenu sur un ensemble de test, c'est-à-dire que le classificateur est utilisé comme une boîte noire pour évaluer les sous-ensembles. Bien que ces techniques puissent atteindre un bon résultat, elles sont coûteuses en termes de calcul. Les méthodes "Filters" [104] supposent une indépendance totale entre le classificateur et les paramètres. Elles utilisent donc une mesure indépendante de l'algorithme de classification pour évaluer les sous-ensembles de paramètres. Les méthodes "Filters" sont relativement robustes, mais peuvent ne pas sélectionner le meilleur sous-ensemble de paramètres pour la classification. Les méthodes hybrides [98] incorporent des connaissances sur la structure spécifique de la classe donnée par le classificateur [105]. Ces méthodes sont généralement moins coûteuses en termes de calcul en comparaison avec les méthodes "Wrappers". En outre, elles sont beaucoup plus lentes que les approches "Filters". Dans ces méthodes, les paramètres sélectionnés dépendent du classificateur.

La génération des sous-ensembles de paramètres, pour les méthodes "Filters" multivariées, dépend de la stratégie de recherche. Bien qu'il existe de nombreuses stratégies de recherche, quatre types sont les plus habituels pour la génération de sous-ensembles de paramètres et sont donnés comme suit :

- Sélection vers l'avant ("Forward");
- Elimination en arrière ("Backward");
- Sélection bidirectionnelle ("Stepwise");
- Sélection heuristique de sous-ensemble de paramètres.

La sélection "Forward" commence généralement par un ensemble vide de paramètres, puis à chaque itération on ajoute un paramètre à l'ensemble de départ. Dans ce type de—sélection, l'élimination commence généralement par l'ensemble des paramètres initial, puis à chaque itération on supprime un paramètre de l'ensemble. La recherche bidirectionnelle "Stepwise" commence des deux côtés en ajoutant un

paramètre pertinent et également en supprimant à la fois un paramètre non pertinent à partir de l'ensemble de départ.

Les stratégies de recherche les plus courantes peuvent être classées en algorithmes exponentiels, algorithmes séquentiels et algorithmes aléatoires. Les algorithmes exponentiels évaluent un certain nombre de sous-ensembles qui croît de façon exponentielle en fonction de la taille de l'espace des paramètres. Les algorithmes séquentiels ajoutent ou suppriment des paramètres de manière séquentielle, ce qui peut conduire à des minimums locaux. Les algorithmes aléatoires incorporent l'aléatoire dans leur procédure de recherche, ce qui évite les minimums locaux [106].

III-3-1 Méthodes de sélection "*Filters*"

Les méthodes "*Filters*" sélectionnent les paramètres en fonction d'un critère de sélection indépendant de l'algorithme de modélisation des données utilisées. Ce n'est qu'après avoir trouvé les meilleurs paramètres que les algorithmes de modélisation peuvent les utiliser. Généralement, nous pouvons classer les mesures développées pour ces méthodes en informations, distance, cohérence, similarité et mesures statistiques. Bien qu'il existe de nombreuses méthodes "*Filters*" décrites dans la littérature, une liste de méthodes courantes est donnée dans le Tableau III-1, accompagnée des références appropriées [107].

Les méthodes "*Filters*" univariées classent généralement un seul paramètre, tandis que les méthodes "*Filters*" multivariées évaluent un sous-ensemble de paramètres.

Dans le cadre de notre thèse de Doctorat, on s'intéresse à la sélection des paramètres statistiques pertinents des harmoniques du courant électrique pour l'identification des appareils électriques en se basant sur l'information mutuelle (IM).

III-3-2 Méthodes "*Wrappers*"

Les méthodes de sélection "*Wrappers*" évaluent la pertinence des sous-ensembles de paramètres par la qualité des performances de l'algorithme de modélisation, qui est considéré comme un évaluateur du sous-ensemble de paramètres à sélectionner. Ainsi, pour les tâches de classification, cette méthode de sélection évaluera les sous-ensembles en fonction des performances du classificateur [108], [109], [110].

Tableau III-1 Différentes méthodes de sélection "Filters" [111]

Noms	Classe de filtre	Applicable à la tâche	Etudier par
Gain d'information	Information Univariée	Classification	[112]
Rapport de gain	Information Univariée	Classification	[113]
Incertitude symétrique	Information Univariée	Classification	[114]
Corrélation	Statistique Univariée		[114]
Chi - carré	Statistique Univariée	Classification	[113]
Critère d'incohérence	Cohérence Multivariée	Classification	[115]
Redondance minimale, pertinence maximale (mRmR)	Information Multivariée	Classification Régression	[116]
Sélection des paramètres basée sur la corrélation (CFS)	Statistique Multivariée	Classification Régression	[113]
Filtrante basée sur la corrélation rapide (FCBF)	Information Multivariée	Classification	[114]
Score de Fisher	Statistique Multivariée	Classification	[117]
Releif et ReleifF	Distance Univariée	Classification Régression	[118]
Sélection des paramètres spectrales (SPEC) et score Laplacien (LS)	Similarité Uni-variée	Classification Regroupement	[97]
Sélection de paramètres pour regroupement	Similarité Multivariée	Regroupement	[119]
Sélection de paramètres localisés basée sur la séparabilité des diffusions (LFSBSS)	Statistique Multivariée	Regroupement	[120]
Sélection des paramètres multi-groupes (MCFS)	Similarité Multivariée	Regroupement	[97]
Pondération de paramètre K-means	Statistique Multivariée	Regroupement	[121]
ReliefC	Distance Univariée	Regroupement	[122]

L'évaluation est répétée pour chaque sous-ensemble et la génération du sous-ensemble dépend de la stratégie de recherche. Les méthodes de sélection "*Wrappers*" sont beaucoup plus lentes que les méthodes de sélection "*Filters*" car elles cherchent à trouver des sous-ensembles suffisamment pertinents en utilisant la phase d'apprentissage et la phase de test qui exigent un temps d'exécution considérable. Il a été empiriquement prouvé que les méthodes de sélection "*Wrappers*" obtiennent des sous-ensembles avec des meilleures performances en comparaison avec les méthodes "*Filters*" car les sous-ensembles sont évalués à l'aide d'un algorithme de modélisation réel [95].

Dans le cadre de notre thèse de Doctorat, la méthode de sélection "*Wrappers*" est appliquée sur l'espace de paramètres obtenus en utilisant l'analyse par ondelettes, à savoir les descripteurs DWE, LWE et WCC . Nous nous intéressons à étudier la pertinence des paramètres de chaque descripteur en utilisant la sélection de paramètres. Nous choisissons l'approche "*Wrappers*" car l'extraction de paramètres considérée prend un espace de paramètres de faible dimension. Nous avons appliqué la méthode de sélection "*Wrappers*" combinée avec l'algorithme de recherche séquentielle *Forward* [95]. Cet algorithme a été utilisé dans [123], [4] ainsi que dans [40]. Les différentes étapes de l'algorithme de sélection d'un sous-ensemble de paramètres à partir d'un ensemble de N paramètres, sont illustrées sur la Figure III- 2 [123].

III-3-3 Méthodes de sélection hybrides

Les méthodes de sélection hybrides effectuent la sélection de paramètres lors de l'exécution de l'algorithme de modélisation. Ces méthodes sont donc intégrées dans l'algorithme entant que fonctionnalité normale ou étendue. Les méthodes de sélection hybrides courantes incluent les différents types d'algorithmes d'arbre de décision [124], mais aussi d'autres algorithmes (par exemple la régression logistique multinomiale et ses variantes [125]). Certaines méthodes hybrides effectuent une pondération des paramètres basée sur des modèles de régularisation avec des fonctions objectives qui minimisent les erreurs d'ajustement.

Des méthodes hybrides ont été proposées pour combiner les meilleures propriétés des méthodes de sélection "*Filters*" et les méthodes de sélection "*Wrappers*". Tout

d'abord, une méthode de type "Filters" est utilisée afin de réduire la dimension de l'espace de paramètres permettant d'obtenir plusieurs sous-ensembles candidats [126]. Ensuite, une sélection de type "wrapper" est utilisée pour trouver le meilleur sous-ensemble candidat. Les méthodes hybrides atteignent généralement une précision élevée qui est une caractéristique des méthodes "Wrappers" et une haute efficacité qui est une caractéristique des méthodes "Filters". Plusieurs méthodes hybrides intéressantes ont été récemment proposées, telles que les algorithmes génétiques hybrides [127], l'optimisation hybride [128], ou l'algorithme de recherche gravitationnelle mixte [129].

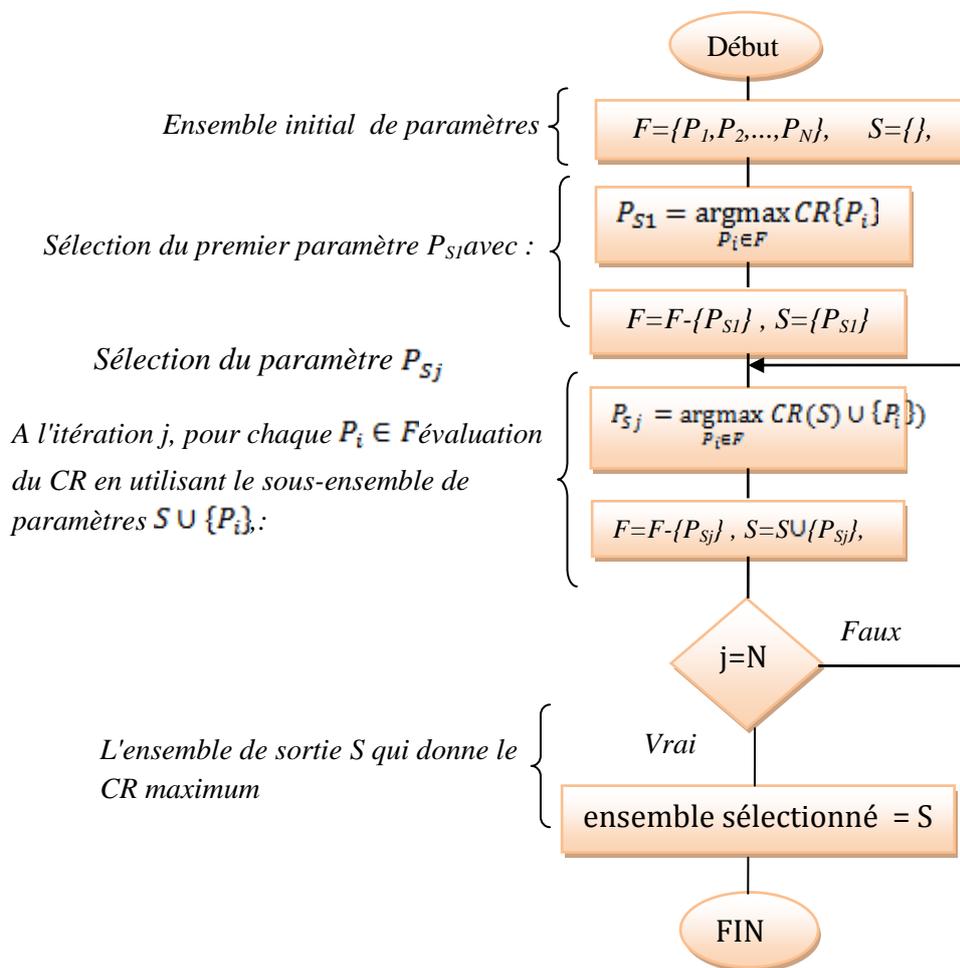


Figure III-2 Algorithme de sélection Wrapper [123].

III-4 Méthodes de sélection fondées sur l'information mutuelle (IM)

Dans la littérature scientifique, plusieurs critères ont été proposés pour évaluer les sous-ensembles de paramètres sélectionnés. Parmi ces critères, on peut citer le taux

d'incohérence [130], la corrélation [131], l'erreur de classification [132], la dimension fractale [133], la mesure de distance [134], etc. L'information mutuelle (IM) est une mesure de la dépendance statistique qui présente deux propriétés principales. Premièrement, elle peut mesurer tout type de relation entre des variables aléatoires, y compris les relations non linéaires [135]. Deuxièmement, l'IM est invariante sous l'effet des transformations de l'espace de paramètres qui sont inversibles et différenciables [136], [137]. Battiti [138] est le premier qui a défini le problème de la sélection des paramètres comme processus de sélection des k paramètres les plus pertinents à partir d'un ensemble initial de m paramètres, tel que $k < m$. Battiti a proposé la sélection séquentielle "forward" d'un seul paramètre à la fois, comme alternative pour évaluer les combinaisons de tous les sous-ensembles de paramètres appartenant à l'ensemble initial. Pour se faire, il adopte l'algorithme de sélection *greedy*. Les principales hypothèses du travail de Battiti sont les suivantes :

- Les paramètres sont classés comme pertinents et redondants ;
- Une fonction heuristique est utilisée pour sélectionner les paramètres, ce qui permet de contrôler le compromis entre la pertinence et la redondance ;
- Le sous-ensemble de paramètres sélectionné est supposé optimal.

L'algorithme de sélection de paramètres *greedy*, fondé sur l'IM, est donné comme suit:

Début d'algorithme

1. $F \leftarrow$ ensemble initial de m paramètres et $S \leftarrow \{\}$ (vide) ;
2. calculer $I(C; Y_i) / \forall Y_i \in F$
3. sélectionner le premier paramètre Y_{p_1} qui maximise $I(C; Y_i)$;
4. $F \leftarrow F - \{Y_{p_1}\}$ et $S \leftarrow S + \{Y_{p_1}\}$;
5. répéter K fois (k est le nombre de paramètres à sélectionner) ;
 - calculer $I(C; S, Y_i) / \forall Y_i \in F$,
 - sélectionner le paramètre Y_{p_j} qui maximise $I(C; S, Y_i)$, à l'étape j
 - $F \leftarrow F - \{Y_{p_j}\}$ et $S \leftarrow S \cup \{Y_{p_j}\}$.

6. faire sortir S .

Fin de l'algorithme

L'objectif principal de cet algorithme consiste donc à trouver une nouvelle définition de la pertinence, ainsi de nouvelles fonctions théoriques d'évaluation avec des stratégies de recherche efficaces menant à des sous-ensembles de paramètres optimaux. Dans les sections suivantes, nous présentons une revue sur les méthodes de sélection des paramètres "Filters" basées sur l'IM. Cette dernière consiste à estimer directement l'IM multivariée entre la classe, les paramètres déjà sélectionnés et le paramètre candidat à sélectionner. Cette estimation dépend des fonctions de densité de probabilités (fdp) conjointe. Cependant, cette mesure peut être compliquée car, à l'exception des faibles dimensions, l'estimation de la fdp de grande dimension peut être très biaisée en utilisant un ensemble de données limité. Ainsi, la plupart des algorithmes de sélection heuristique utilisent au maximum trois paramètres (variable de classe, paramètre sélectionné et paramètre à sélectionner) [93], [45]. Dans ce qui suit, nous allons présenter des notions sur la théorie de l'information, notamment l'estimation de l'information mutuelle et l'entropie, ainsi que les différentes stratégies heuristiques de sélection de paramètres basées sur le critère de maximisation de l'information mutuelle. Parmi celles-ci, les méthodes les plus populaires sont la CIFE [139], la JMI [140], la CMIM [141] et la JMIM [142].

III-4-1- Entropie et Information mutuelle

L'entropie et l'information mutuelle sont des concepts de base définis dans la théorie de l'information. À son origine, la théorie de l'information était utilisée dans le contexte de la théorie de la communication, pour trouver des réponses sur la compression des données et le taux de transmission. Depuis, les principes de la théorie de l'information ont été largement intégrés dans l'apprentissage automatique [143].

III-4-1-1- Entropie (information moyenne)

L'entropie est définie par Shannon comme une mesure d'incertitude d'une variable aléatoire. L'entropie est définie aussi comme la plus petite quantité moyenne de bits capable de représenter une information tirée de la variable aléatoire [144].

Soit X une variable aléatoire discrète. L'entropie $H(X)$ de X est définie par :

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log_2(p(x)) \quad (III - 1)$$

où $p(x) = Pr(X = x)$ est la fonction de densité de probabilité de X . Notez que l'entropie ne dépend pas des valeurs réelles, juste de la distribution de probabilité de la variable aléatoire.

L'entropie conjointe

Pour deux variables aléatoires discrètes X et Y caractérisés par leur fonction de densité de probabilité $p(x, y)$, l'entropie conjointe $H(X, Y)$ est définie comme:

$$H(\{x, y\}) = H(X, Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2(p(x, y)) \quad (III - 2)$$

L'entropie conjointe a des valeurs dans la gamme définie par l'inégalité suivante:

$$\max(H(x), H(y)) \leq H(\{x, y\}) \leq H(x) + H(y) \quad (III - 3)$$

La valeur maximale de l'inégalité (III-3) se produit lorsque x et y sont complètement indépendants. La valeur minimale se produit lorsque x dépend complètement de y .

L'entropie conditionnelle

L'entropie conjointe mesure l'incertitude restante de la variable aléatoire x lorsque la valeur de la variable aléatoire y est connue. La valeur minimale de l'entropie conditionnelle est nulle lorsque x dépend statistiquement de y . La valeur maximale se produit lorsque x et y sont statistiquement indépendants. L'entropie conditionnelle $H(X \setminus Y)$ de X par rapport à Y est donnée par:

$$H(X \setminus Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2(p(x|y)) \quad (III - 4)$$

où $0 < H(X \setminus Y) < H(X)$

Une autre façon de représenter l'entropie conditionnelle est:

$$H(X|Y) = H(X, Y) - H(Y) \quad (III - 5)$$

Dans le cas où les deux variables aléatoires X et Y sont des variables continues, $h(X)$, $h(X, Y)$ et $h(X|Y)$ sont données par :

$$h(X) = - \int_x f_X(x) \ln f_X(x) dx \quad (III - 6)$$

$$h(X, Y) = - \int_y \int_x f_{XY}(x, y) \ln f_{XY}(x, y) dx dy \quad (III - 7)$$

$$h(X|Y) = - \int_y \int_x f_{XY}(x, y) \ln f_{X|Y=y}(x) dx dy \quad (III - 8)$$

III-4-1-2- Information mutuelle

L'IM est une mesure de la distance entre deux distributions de probabilité. Entre deux variables discrètes X et Y, l'IM $I(X; Y)$ est définie en utilisant leur distribution de probabilité conjointe $p(x, y)$ et leurs probabilités marginales $p(x)$ et $p(y)$. Elle est donnée par: [145], [146], [147], [148], [141].

$$I(X; Y) = H(X) - H(X|Y) \quad (III - 9)$$

$$= H(Y) - H(Y|X) \quad (III - 10)$$

$$= H(X) + H(Y) - H(X, Y) \quad (III - 11)$$

$$= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (III - 12)$$

où $I(X; Y) = 0$ lorsque X et Y sont statistiquement indépendantes, c'est-à-dire $p(x, y) = p(x) \cdot p(y)$. L'IM est liée linéairement aux entropies des variables à travers l'équation suivante :

$$I(x; y) = \begin{cases} H(x) - H(x|y) \\ H(y) - H(y|x) \\ H(x) + H(y) - H(x, y) \end{cases} \quad (III - 13)$$

Cette définition de l'information mutuelle est utile dans le contexte de la sélection des paramètres car elle permet de quantifier la pertinence d'un sous-ensemble de paramètres par rapport au vecteur de sortie C (Class). La Figure III-3 présente un diagramme de relation entre l'entropie et l'IM

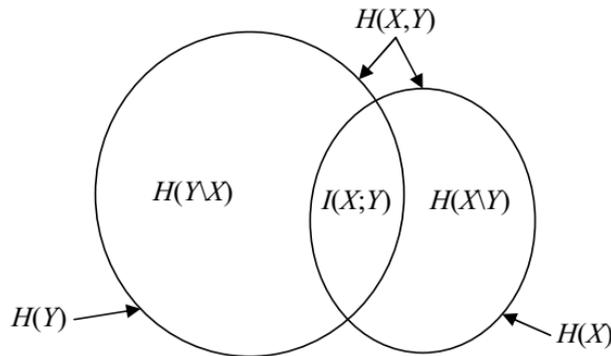


Figure III-3 Diagramme reliant l'IM et l'entropie [145]

III-4-1-3 Information mutuelle d'ordre supérieur

Lorsqu'il s'agit de plus de deux variables, il apparaît la nécessité de calculer l'IM entre trois variables ou plus. L'une des principales définitions de l'IM entre-trois variables est l'IM triple [149]. L'IM entre les variables aléatoires continues X, Y, Z, ayants la fonction de densité de probabilité conjointe $f(x; y; z)$, et les distributions marginales $f_{XY}(x; y), f_{XZ}(x; z), f_{YZ}(y; z), f_X(x), f_Y(y)$ et $f_Z(z)$ est donnée par :

$$I(X; Y; Z) = \int_z \int_y \int_x f(x, y, z) \ln \frac{f_{XY}(x, y) f_{XZ}(x, z) f_{YZ}(y, z)}{f(x, y, z) f_X(x) f_Y(y) f_Z(z)} dx dy dz \quad (III - 14)$$

En utilisant cette définition, nous pouvons démontrer que, pour les variables aléatoires X et Y et une variable aléatoire ou un vecteur aléatoire Z, nous avons :

$$\begin{aligned} I(X; Y; Z) &= I(X; Y) - I(X; Y \setminus Z) \\ &= I(X; Z) - I(X; Z \setminus Y) \\ &= I(Y; Z) - I(Y; Z \setminus X) \end{aligned} \quad (III - 15)$$

où

$$I(X; Y|Z) = h(X|Z) - h(X|Y, Z)$$

L'IMI $I(X; Y; Z)$ peut avoir des valeurs positives ou négatives [149]. En effet, l'IMI $I(X; Y; Z)$ ayant une valeur positive signifie que X et Y sont des variables aléatoires redondantes pour expliquer la variable aléatoire Z. Dans le cas contraire, X et Y sont des variables aléatoires synergiques.

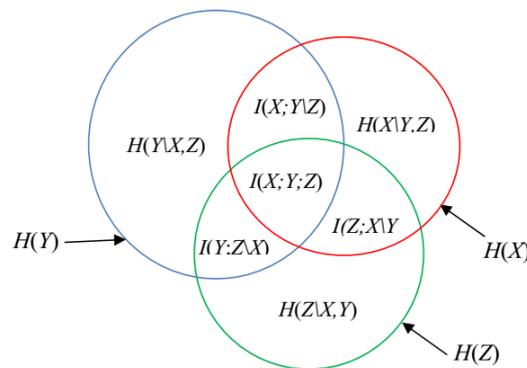


Figure III-4 Information mutuelle multivariée de trois variables [150]

L'IM est symétrique, c'est-à-dire $I(x; y; z) = IM(x; z; y) = I(z; y; x) = I(y; x; z) \dots$ L'IM n'a pas été largement utilisée dans la littérature, en raison de son interprétation difficile du fait qu'elle peut prendre par exemple, des valeurs négatives. Cependant, il existe des articles intéressants, travaillant sur l'interaction entre les variables, qui utilisent ce concept [151], [152], [153], [154]. L'IM peut être définie comme étant la quantité d'informations communes à toutes les variables (ou ensemble de variables), mais qui n'est présente dans aucun sous-ensemble de ces variables.

III-4-2 Estimation de l'entropie et de l'information mutuelle

Le calcul de l'IM provient de l'estimation des fonctions fdp conjointes et marginales. En pratique, ces fonctions ne sont pas connues et elles doivent être estimées à partir d'un nombre fini d'échantillons. Plusieurs méthodes d'estimation de l'entropie et de l'information mutuelle sont proposées telles que la méthode d'histogramme [155], la méthode à noyau [156], ou la méthode de mélange Gaussien GMM (Gaussian mixture model) [157]. Dans notre travail, on s'intéresse à l'estimation par la méthode d'histogramme vue sa simplicité [45]. Cependant, cette méthode souffre du problème du choix du nombre k de *bins* (cellules) de discrétisation de l'histogramme. Il existe

plusieurs critères heuristiques pour le choix de la valeur de k qui prennent en considération le nombre d'observations et la fdp de référence Gaussienne. On peut citer comme exemples les méthodes Sturges [158], Scott [159], Freedman [160] et HGR [93].

III-4-2-1 Formules du choix optimal du nombre de bins

Soit X une variable aléatoire, le nombre de bins k peut être donné par:

$$k = \frac{A_x}{h} \tag{III - 16}$$

où $A_x = x_{max} - x_{min}$ est l'étendue de la variable x et h est la largeur de bins.

Les formules suivantes peuvent également donner le nombre de bins k ou la largeur de bins h pour une donnée de N échantillons et σ l'écart type de la variable X

- Formule de sturges [158]:

$$k = 1 + \log_2(N) \tag{III - 17}$$

- Formule de Scott [159]:

$$h = \frac{3.5\sigma}{N^{(1/5)}} \tag{III - 18}$$

- Formule de Freedman-Diaconis [160]:

$$h = 2 \cdot \frac{IQR(x)}{N^{(1/3)}} \tag{III - 19}$$

où IQR est l'intervalle interquartile.

- Formule de HGR [93]:

➤ Le nombre de bins k proposé par [93] pour estimer l'entropie marginal $H(X)$ est donné par:

$$k_{opt} = \text{round} \left\{ \frac{\zeta}{6} + \frac{2}{3\zeta} + \frac{1}{3} \right\} \tag{III - 20}$$

Avec

$$\zeta = \sqrt[3]{(8 + 108G + 12^2 \sqrt{12G + 81G^2})} \tag{III - 21}$$

$$G = \frac{N A_x^2}{12 \hat{\sigma}_x^2} \tag{III - 22}$$

où

$\hat{\sigma}_x^2$: est l'estimée pratique de l'écart type

- Le nombre de *bins* k pour estimer l'information mutuelle par la méthode d'histogramme conjoint est donné par:

$$k = \text{round} \left\{ \frac{1}{2} + \frac{1}{2} \sqrt{1 + 4\sqrt{L}} \right\} \quad (\text{III} - 23)$$

$$\text{avec} \quad L = \frac{N\rho^2}{12(1-\rho^2)} (\alpha_x^2 + \alpha_y^2) \quad (\text{III} - 24)$$

où

α_x, α_y sont des constants ;

ρ est le coefficient de corrélation ;

Dans le cas d'une distribution gaussienne, le nombre de *bins* k pour l'estimateur de l'IM basé sur l'histogramme est donné par :

$$k = \text{round} \left\{ \frac{1}{2} + \frac{1}{2} \sqrt{1 + 4 \sqrt{\frac{6N\hat{\rho}^2}{1 - \hat{\rho}^2}}} \right\} \quad (\text{III} - 25)$$

où $\hat{\rho}$ est l'estimateur de coefficient de corrélation ρ .

Dans notre travail, nous avons utilisé les formules proposées par [93].

III-4-3 Stratégies de sélection "*Filters*" basée sur l'IM

Les méthodes de sélection des paramètres partagent l'objectif général d'identifier un sous-ensemble approprié S de k paramètres $\{Y_{P_1}, Y_{P_2}, \dots, Y_{P_k}\}$ à partir des paramètres originaux F de n paramètres $\{Y_1, Y_2, \dots, Y_n\}$, avec la propriété d'être au maximum informative sur la variable de classe C pour la tâche de classification [161], [146]. Nous appelons le sous-ensemble de paramètres pertinents, de taille minimale, le sous ensemble S_{opt} [147]. La quantité d'informations apportée par S_{opt} est souvent évaluée grâce à l'information mutuelle (IM) comme suit:

$$S_{\text{opt}} = \arg \max_{S \in F} I(C; S) \quad (\text{III} - 26)$$

La sélection du sous-ensemble S_{opt} par l'algorithme séquentiel "*greedy Forward*" consiste à sélectionner, pour chaque itération j , un paramètre Y_{P_j} qui maximise

l'information mutuelle $I(C; S, Y_{p_j})$ entre la variable classe C , le sous ensemble de paramètres déjà sélectionné à l'itération $j-1$ et le paramètre à sélectionner Y_{p_j} [146]

$$\begin{aligned}
 Y_{p_j} &= \arg \max_{Y_i \in F - S_{j-1}} I(C; S_{j-1}, Y_i) \\
 &= \arg \max_{Y_i \in F - S_{j-1}} [I(C; S_{j-1}) + I(C; Y_i \setminus S_{j-1})] \\
 &= \arg \max_{Y_i \in F - S_{j-1}} [I(C; Y_i \setminus S_{j-1})] \\
 Y_{p_j} &= \arg \max_{Y_i \in F - S_{j-1}} [I(C; Y_i) - I_3(C; Y_i; S_{j-1})] \quad (\text{III} - 27)
 \end{aligned}$$

Statistiquement, il n'est pas facile de calculer la valeur de $I_3(C; Y_i; S_{j-1})$ lorsque j prend de grandes valeurs. A cet effet, plusieurs méthodes heuristiques ont été proposées telles que la MIM, la CMIM, la MIFS, la MRMR, la JMI. Dans ces méthodes, le nombre de variables aléatoires est limité à trois (la variable classe C et deux autres paramètres) [146]. Dans ce qui suit une perf présentation de ces méthodes.

- MIM (Mutual Information Maximization) [162]: cette stratégie est la plus simple, car elle ne tient en compte que les informations mutuelles individuelles de chaque paramètre avec la variable C . Ainsi, cette stratégie ne considère pas la redondance avec les paramètres déjà sélectionnés, ce qui conduit à sélectionner des paramètres redondants [163]. La sélection du paramètre Y_{p_j} à l'itération j s'effectue selon la formule suivante:

$$Y_{p_j}(\text{MIM}) = \arg \max_{Y_i \in F - S_{j-1}} [I(C; Y_i)] \quad (\text{III} - 28)$$

- CMIM (Conditional Mutual Information Maximization) [141]: Cette méthode sélectionne de manière itérative un paramètre qui maximise son IM conditionnelle avec la variable de classe C , et chaque paramètre déjà sélectionné. La sélection du paramètre Y_{p_j} à l'itération j s'effectue selon la formule suivante:

$$Y_{p_j}(\text{CMIM}) = \arg \max_{Y_i \in F - S_{j-1}} \left[\min_{Y_k \in S} [I(C; Y_i \setminus Y_k)] \right] \quad (\text{III} - 29)$$

$$= \arg \max_{Y_i \in F-S_{j-1}} \left[\min_{Y_k \in S} [I(C; Y_i) - I(C; Y_i; Y_k)] \right] \quad (III - 30)$$

$$= \arg \max_{Y_i \in F-S_{j-1}} \left[I(C; Y_i) - \max_{Y_k \in S} [I(C; Y_i; Y_k)] \right] \quad (III - 31)$$

$$= \arg \max_{Y_i \in F-S_{j-1}} \left[I(C; Y_i) - \max_{Y_k \in S} [I(Y_i; Y_k) - I(Y_i; Y_k \setminus C)] \right] \quad (III - 32)$$

- MIFS (Mutual Information Feature Selection) [138]: Cette méthode ne tient en compte que la IM entre deux variables aléatoires. Elle suit la formule suivante:

$$Y_{p_j}(\text{MIFS}) = I(Y_i; C) - \beta \sum_{Y_i \in S} I(Y_i; Y_j) \quad (III - 33)$$

- MRMR(Max-Relevance Min-Redundancy) [164]: Dans cette méthode l'estimation de la IM est basée sur la maximisation de la pertinence et la minimisation de la redondance entre deux variable aléatoire. La sélection du paramètre Y_{p_j} à l'itération j s'effectue selon la formule suivante:

$$Y_{p_j}(\text{MRMR}) = \arg \max_{Y_i \in F-S_{j-1}} \left[I(C; Y_i) - \frac{1}{j-1} \sum_{k=1}^{j-1} I(Y_i; Y_{p_k}) \right] \quad (III - 34)$$

- JMI (Joint Mutual Information) [140] proposent l'utilisation de l'information mutuelle conjointe JMI. Dans cette stratégie, la sélection du paramètre Y_{p_j} à l'itération j s'effectue selon la formule suivante:

$$Y_{p_j}(\text{JMI}) = \arg \max_{Y_i \in F-S_{j-1}} \left[I(C; Y_i) - \frac{1}{j-1} \sum_{k=1}^{j-1} I(C; Y_i; Y_{p_k}) \right] \quad (III - 35)$$

III-5 Application de la sélection des paramètres pertinents dan les systèmes IAED

Dans le chapitre II, deux méthodes d'extraction de paramètres statistiques ont été proposées pour la conception des systèmes IAED. La première méthode d'extraction basée sur l'analyse harmonique, consiste à convertir chaque signal du courant électrique en une séquence de vecteurs constitué chacun de 1000 paramètres statistiques. Ainsi, cette méthode permet d'extraire des vecteurs de grandes dimensions qui peuvent

causer probablement le phénomène de la malédiction et la détérioration des performances du système d'identification en termes de précision et de complexité. Dans le but d'éviter ce phénomène, on propose d'appliquer des méthodes de sélection de type "Filters" basées sur le critère de maximisation de l'information mutuelle pour réduire la dimensionnalité de vecteurs de paramètres [5]. La deuxième méthode consiste à extraire des vecteurs de paramètres statistiques en se basant sur l'analyse par ondelette. Cette méthode permet d'obtenir une représentation plus compacte des vecteurs de paramètres de faibles dimensions par rapport à celle de l'analyse harmonique. Dans le but d'étudier la pertinence des ces paramètres, on applique des méthodes de sélection de type "Wrappers" [7].

III-6 Conclusion

Dans ce chapitre, nous avons présenté un état de l'art sur les méthodes de sélection de paramètres pertinents et leurs étapes de fonctionnement. Les approches de sélection couramment utilisées sont l'approche "Wrappers" et l'approche Filters. La première approche, utilisant le critère d'évaluation des sous-ensembles basé sur la précision de classificateur, est plus adaptée pour les espaces de paramètres à faible dimension. Alors que la deuxième approche, indépendante du classificateur, est utilisée pour le cas des espaces de paramètres de grandes dimensions. Ces deux approches sont couramment appliquées avec la génération séquentielle des sous-ensembles en utilisant l'algorithme "greedy Forward".

Dans notre travail, deux propositions de sélection de paramètres ont été décrites de manière abrégée. La première consiste à appliquer la méthode de sélection de type *Filters* basée sur le critère de maximisation de l'information mutuelle, sur l'ensemble de paramètres harmoniques pour la tâche IAED. La deuxième proposition consiste à appliquer la méthode de sélection de types "Wrappers", sur l'ensemble de paramètres extraits de l'analyse par ondelettes. Ces propositions seront discutées en détails dans le chapitre IV.

CHAPITRE IV:
SYSTEME PROPOSE:
IMPLEMENTATION ET RESULTATS

IV-1 Introduction

Un système IAED est une application des systèmes de reconnaissance de formes, qui consiste à identifier la classe de l'appareil électrique à partir du signal électrique. Le fonctionnement de tel système se base principalement sur une étape d'extraction de paramètres et une étape d'identification en utilisant un classificateur tel que KNN, SVM, GMM,...etc. La conception d'un système IAED exige une phase d'apprentissage pour modéliser les différentes classes et une phase de test pour évaluer les performances du système IAED. Cette évaluation s'effectue en utilisant une base de données (BD) répartie en une BD d'apprentissage et une BD de test.

Plusieurs systèmes IAED proposés sont basés sur les modèles HMM [1], [40]. Ces systèmes permettent de classifier toute la séquence de vecteurs de paramètres obtenue par la méthode d'extraction considérée, en une classe d'appareil en utilisant le classificateur HMM. Dans notre travail, on s'intéresse à concevoir un système IAED basé sur le classificateur KNN pour minimiser la complexité et probablement améliorer la précision du système. Un tel système consiste à classifier chaque vecteur de la séquence de vecteurs de paramètres, puis à classifier la séquence des indices de classes obtenue par l'application de la méthode de règle de vote pour décider sur la classe du signal d'entrée. Cependant, un grand nombre de signaux de la base d'apprentissage peut provoquer l'augmentation du temps de calcul et de l'espace mémoire. L'objectif principal de notre travail est de minimiser cette complexité, en minimisant le nombre initial des vecteurs de la base d'apprentissage, et également en réduisant la dimensionnalité des vecteurs de paramètres. La première solution de minimisation consiste à proposer une nouvelle méthode d'extraction de paramètres statistiques décrite dans la section II-3-3. Cette méthode est appliquée sur les séquences de paramètres harmoniques (voir section II-4). La deuxième solution consiste à réduire la dimensionnalité, soit en cherchant une représentation plus compacte basée sur l'analyse par ondelettes, soit en sélectionnant les paramètres les plus pertinents par l'application des algorithmes de sélection.

Le premier système IAED à proposer se base sur l'extraction et la sélection de paramètres statistiques des harmoniques. La sélection de paramètres s'effectue en appliquant l'approche de sélection de type "Filters" basée sur la maximisation de l'information mutuelle [5]. Le deuxième système IAED se base sur l'extraction et la

sélection de paramètres statistiques des descripteurs DWE, LDW et WCC obtenues en utilisant l'analyse par ondelettes. La sélection des paramètres s'effectue en appliquant l'approche "Wrappers" [6] [7]. Les deux systèmes IAED utilisent le classificateur KNN combiné avec la stratégie de la règle de vote.

Plusieurs expériences seront menées pour chercher les bonnes configurations des systèmes proposés. La configuration du premier système consiste à chercher:

- La meilleure combinaison des moments statistiques (moyenne, écart-type, l'asymétrie et l'aplatissement);
- Le nombre optimal de vecteurs statistiques par seconde;
- Les paramètres les plus pertinents par l'application de la méthode de sélection de paramètres basée sur la maximisation de l'information mutuelle.

La configuration du deuxième système consiste à chercher:

- Le nombre optimal de vecteurs statistiques par seconde pour les différents descripteurs (DWE, LWE et WCC);
- La meilleure durée de fenêtre d'analyse;
- La configuration optimale pour la famille d'ondelettes et le niveau de décomposition;
- La pertinence des paramètres en utilisant la méthode de sélection "Wrappers".

Les deux systèmes utilisent la même architecture mais ils se diffèrent au niveau des étapes d'extraction et de sélection de paramètres. Ainsi, dans la section suivante on décrira l'architecture globale des systèmes proposés.

IV-2 Architecture globale des systèmes proposés

L'architecture générale des systèmes proposés est illustrée sur la Figure IV-1. Deux types d'analyses sont envisagés à savoir L'analyse harmonique et l'analyse par ondelettes. Plus particulièrement, l'analyse harmonique est combinée avec la méthode de sélection de type "*Filters*", alors que l'analyse par ondelettes est combinée avec la méthode de sélection de type "*Wrappers*". Ainsi, cette architecture nous permet de concevoir deux systèmes IAED basés sur le classificateur KNN combiné avec la stratégie de règle de vote [5] [6] [7].

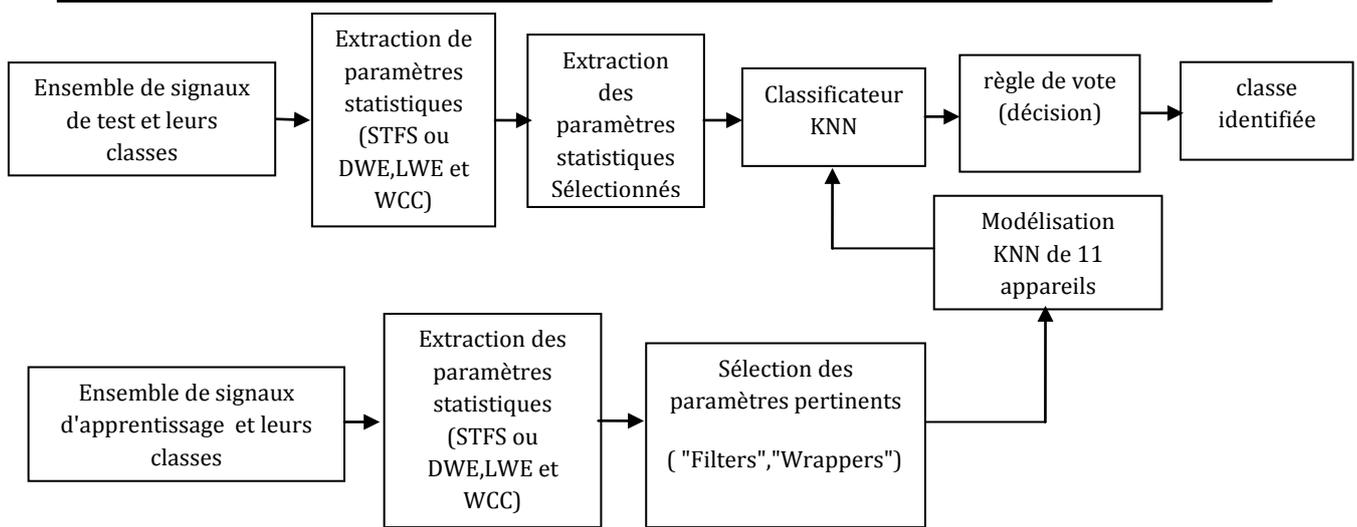


Figure IV-1: Architecture globale des Systèmes IAED proposés [[HYPERLINK \l](#)]

IV-2-1 Base de donnée PLAID

L'évaluation des performances des systèmes proposés se base sur la base de données PLAID [165] qui se compose de 1074 enregistrements du courant et de la tension de 11 types d'appareils électriques provenant des installations de 55 maisons. Les enregistrements sont réalisés à une fréquence de 30 kHz. Le tableau IV-1 présente un récapitulatif sur la base de données PLAID et ses caractéristiques.

Tableau IV-1: Récapitulatif sur la base de données PLAID

N	Type d'appareil	Nombre d'instances
1	Lampe fluorescente compacte	175
2	Aspirateur	38
3	Sèche-cheveux	156
4	Micro-onde	139
5	Climatiseur	66
6	PC Portable	172
7	Réfrigérateur	38
8	Ampoule à lumière incandescence	114
9	Ventilateur	115
10	Machine à laver	26
11	Chauffe-eau	35
	Globalement	1074

IV-2-2 Préparation de données

L'évaluation des performances des systèmes IAED exige de répartir la base de données en une base d'apprentissage et une base de test. Dans notre travail, la base de données PLAID est répartie initialement en une base d'apprentissage de 537 signaux (50%) et une base de données de test de 537 signaux (50%) du courant électrique. Cette répartition tient en compte le fonctionnement du système en mode dépendant de la maison (house) qui se caractérise par l'appartenance de chaque maison avec des exemples dans la base d'apprentissage et dans la base de test. L'effet de la répartition sur les performances des systèmes sera étudié ultérieurement. Le tableau IV-2 illustre la répartition initiale de la base de données PLAID en détaillant le nombre d'échantillons pour chaque appareil électrique.

Tableau IV-2: Nombre d'instances pour une répartition 50%/50% de la base de données PLAID.

N	Type d'appareil	Nombre d'instances dans la base d'apprentissage	Nombre d'instances dans la base de test
1	Lampe fluorescente compacte	87	88
2	Aspirateur	24	14
3	Sèche-cheveux	74	82
4	Micro-onde	69	70
5	Climatiseur	33	33
6	PC Portable	86	86
7	Réfrigérateur	19	19
8	Ampoule à lumière incandescence	57	57
9	Ventilateur	57	58
10	Machine à laver	13	13
11	Chauffe-eau	18	17
	Globalement	537	537

IV-2-3 Extraction des paramètres statistiques

L'extraction des paramètres statistiques consiste à convertir chaque signal de courant électrique en une séquence de vecteurs de paramètres. Cette étape est effectuée en suivant plusieurs étapes à savoir l'étape de prétraitement, l'étape de fenêtrage et l'étape d'analyse des signaux utilisant l'une des techniques à savoir l'analyse harmonique ou l'analyse par ondelettes. La Figure IV-2 illustre le schéma synoptique de l'étape d'extraction des paramètres statistiques proposée.

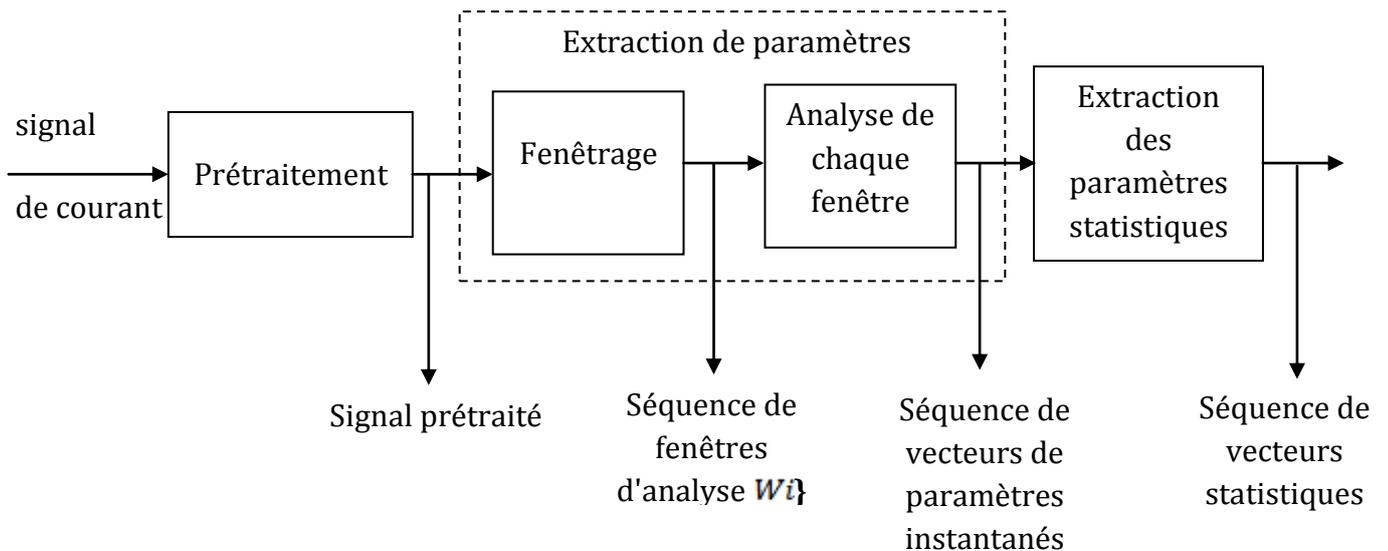


Figure IV-2: Etape d'extraction des paramètres statistiques

IV-2-3-1 Prétraitement

L'objectif de cette étape consiste à chercher l'intervalle de fonctionnement de l'appareil en état de marche. La détection du début et de la fin de l'intervalle de fonctionnement, pour chaque signal de la base PLIAD, a été effectuée dans [1] en utilisant un algorithme de seuillage de l'énergie. Cet algorithme permet également de segmenter cet intervalle en segment transitoire et segment permanent. L'effet de cette segmentation a été discuté dans [1]. Dans notre travail, on ne s'intéresse qu'à la partie entière de l'intervalle de fonctionnement. Ce prétraitement nous permet de supprimer les régions dans lesquelles l'appareil est inactif.

IV-2-3-2 Fenêtrage

L'objectif principal de l'étape de fenêtrage est d'assurer la stationnarité du signal du courant électrique pour une analyse à court terme. Cette étape consiste à découper le

signal du courant électrique en une séquence de fenêtres d'analyse de durée égale à $1/60$ s (16.7 ms) chevauchées de 50%. Chaque fenêtre d'analyse W_i est composée de 500 échantillons (30KHz/60Hz). Le chevauchement de 250 échantillons entre les fenêtres d'analyse a pour objectif d'éviter la perte d'information. Le choix de fenêtre d'analyse de 500 échantillons peut être justifié par la quasi périodicité du signal du courant électrique. Cette étape est utilisée souvent dans les systèmes IAED [1], [40]. La figure IV-3 présente un exemple de fenêtrage d'un signal du courant électrique.

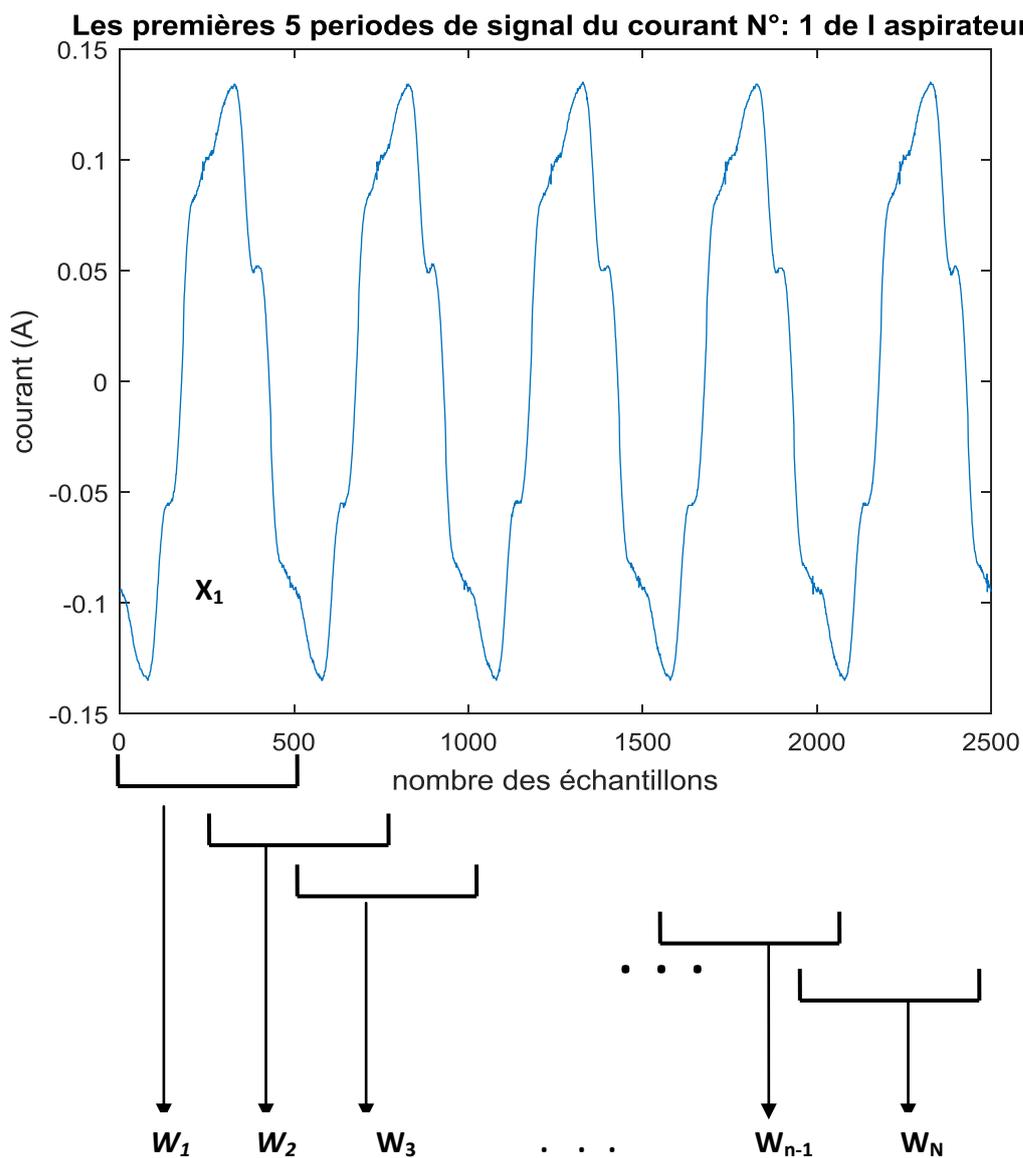


Figure IV-3 Etape de fenêtrage

IV- 2-3-3 Analyse du signal à court terme

Cette étape consiste à convertir chaque fenêtre d'analyse, obtenue dans l'étape précédente, en un vecteur de paramètres. Ainsi, chaque signal du courant électrique sera converti en une séquence de vecteurs de paramètres. Les techniques d'analyse considérées dans notre travail sont l'analyse harmonique destinée à être utilisée avec le premier système IAED proposé et l'analyse par ondelettes destinée à être utilisée avec le deuxième système IAED proposé .

IV- 2-3-3-1 Analyse harmonique à court terme

L'analyse harmonique à court terme consiste à convertir chaque fenêtre du signal échantillonné à 30 KHz en un vecteur de paramètres STFS (voir la section II-4). Puisque la bande fréquentielle maximale du signal est égale à 15 KHz ($F_e/2$), et la fréquence fondamentale est égale à 60 Hz, alors le nombre maximal d'harmoniques est égal à 250 . Ainsi, chaque fenêtre d'analyse sera convertie en un vecteur de 250 paramètres STFS. La figure IV-4 illustre cette analyse.

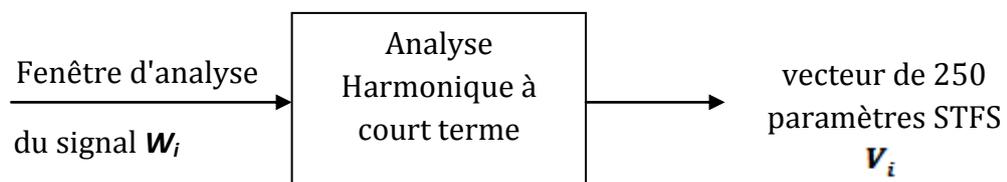


Figure IV-4: Analyse harmonique à court terme de la fenêtre d'analyse

IV- 2-3-3-2 Analyse par ondelettes à court terme

L'analyse par ondelettes à court terme consiste à convertir chaque fenêtre d'analyse en un vecteur de paramètres calculés en appliquant la transformée en ondelettes discrète (DWT). L'extraction de paramètres de chaque fenêtre d'analyse se base sur le calcul de l'énergie à chaque niveau dyadique de la DWT. Le nombre de paramètres obtenu par une analyse de NV décompositions dyadiques est égal à NV+1. Trois descripteurs de paramètres ont été extraits à partir de cette analyse, à savoir: le DWE, le LWE et le WCC (voir section II-5-4). La figure IV-5 illustre l'extraction de paramètres par cette analyse.

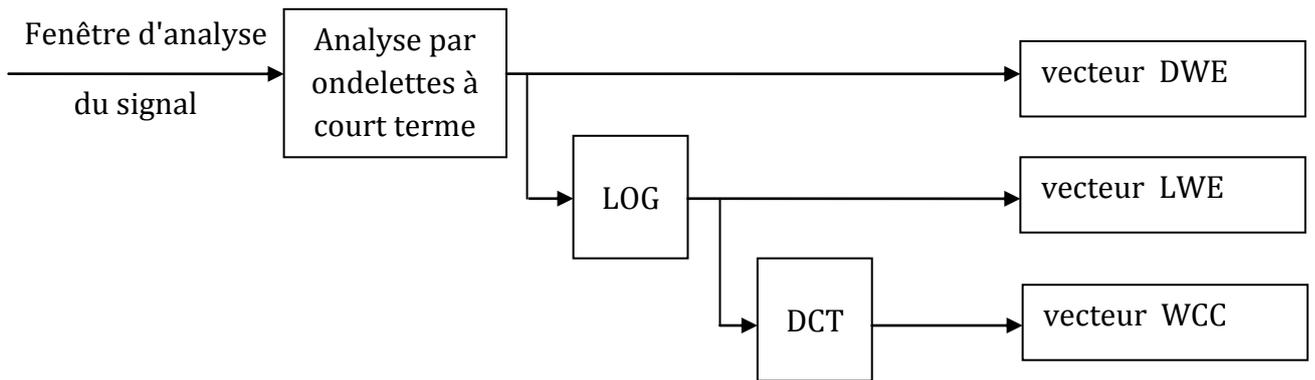


Figure IV-5: Analyse par ondelettes de la fenêtre d'analyse

IV-2-3-4 Estimation des paramètres statistiques

L'originalité de notre approche réside dans le calcul des paramètres statistiques. En effet, elle consiste à appliquer des statistiques sur des sous-séquences de vecteurs de paramètres obtenues dans l'étape d'analyse. Chacune de ces sous-séquences est convertie en un vecteur de paramètres statistiques. Cette étape d'extraction des paramètres statistiques nous permet de minimiser la complexité du système en termes de temps de calcul et d'encombrement mémoire et probablement améliorer les performances du système. Plus particulièrement, elle réduit le nombre de vecteurs de paramètres de la base d'apprentissage et ainsi réduit le nombre d'opérations de comparaison durant l'étape de classification.

La figure suivante illustre la conversion d'une séquence de vecteurs de paramètres $S_v = [V_1, V_2, \dots, V_N]$ en une séquence de vecteurs statistiques $S_{vs} = [V_{s1}, V_{s2}, \dots, V_{sR}]$ durant l'étape d'extraction de paramètres statistiques.

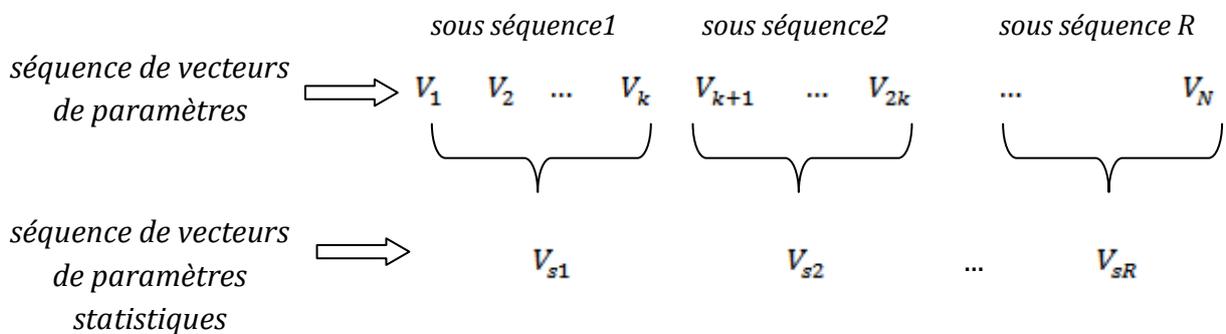
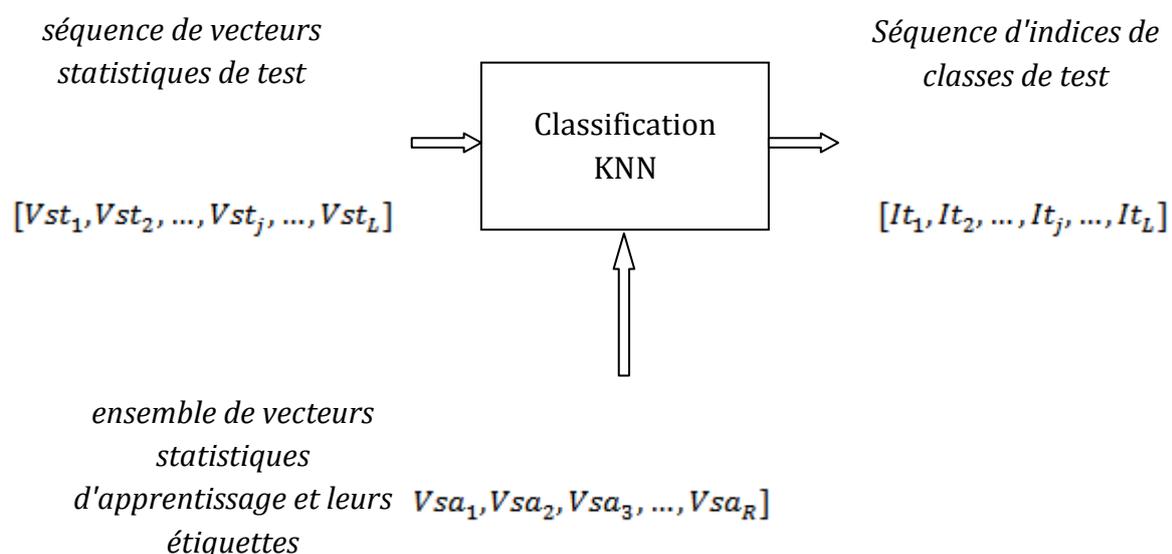


Figure IV-6: Etape d'extraction des vecteurs de paramètres statistiques

Dans le cas de l'analyse harmonique, chaque vecteur de paramètres statistiques est composé de 1000 paramètres statistiques STFS en considérant les statistiques de la moyenne, de l'écart type, de l'asymétrie et de l'aplatissement (voir section II-3-3). Dans le cas d'analyse par ondelettes de NV niveaux de décomposition, chaque vecteur de paramètres statistiques est composé de $4(NV+1)$, en considérant, les mêmes statistiques précédentes.

IV-2-4 Classification KNN

L'étape de classification KNN consiste à classifier chaque vecteur de paramètres statistiques d'une séquence correspondant au signal d'entrée de test, en une classe d'appareil. Cette classification exige préalablement d'enregistrer les séquences de vecteurs de paramètres statistiques correspondant aux signaux d'apprentissage ainsi que leurs étiquettes de classes d'appareils. La classification KNN compare chaque vecteur de test avec l'ensemble des vecteurs d'apprentissage. Puis, elle applique la stratégie de règle de vote sur les K les plus proches voisins du vecteur de test, pour lui attribuer la classe la plus votée. Ainsi, cette classification permet d'obtenir une séquence d'indices de classes des appareils. La figure suivante illustre l'étape de classification KNN. Cette figure présente la classification de chaque vecteur d'une séquence de vecteurs statistiques de test $[Vst_1, Vst_2, \dots, Vst_j, \dots, Vst_L]$ en une séquence d'indices de classes $[It_1, It_2, \dots, It_j, \dots, It_L]$ dont chaque indice appartient à l'ensemble des 11 classes d'appareils.



$$[Ca_1, Ca_2, Ca_3, \dots, Ca_R]$$

D

Figure IV-7: Etape de classification KNN

le but de chercher la classe d'appareil la plus votée à partir de la séquence d'indices de classe obtenue précédemment.

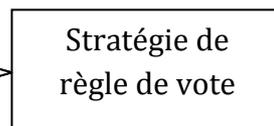
Dans la section suivante, on décrira l'étape de classification du signal de test à partir de la séquence d'indices en utilisant la stratégie de règle de vote.

IV-2-5 Application de la stratégie de règle de vote

La stratégie de règle de vote est issue des sciences sociales [166]. Cette méthode est devenue une technique largement utilisée dans diverses disciplines, notamment dans les disciplines de l'ingénierie, et en particulier dans la reconnaissance de formes [167], [168], [169], [170], [171]. Dans nos systèmes IAED, la procédure de la règle de vote est combinée avec le classificateur KNN pour décider sur la classe d'appareil correspondant au signal d'entrée. Cette approche consiste à appliquer la procédure de règle de vote sur la séquence d'indices de classes pour finalement lui attribuer la classe d'appareil la plus votée. La figure suivante présente l'application de la stratégie de règle de vote sur une séquence d'indices de classes de test correspondant à un signal de test.

Séquence d'indices de classes
de test correspondant à un
signal de test

$$[It_1, It_2, \dots, It_j, \dots, It_L]$$



Classe d'appareil
identifiée

$$C_v$$

Figure IV-8: Application de la stratégie de la règle de vote

IV-2-6 Sélection des paramètres statistiques pertinents

Cette étape consiste à sélectionner les paramètres statistiques les plus pertinents à partir d'un ensemble initial de paramètres statistiques. Dans le premier système IAED basé sur l'analyse harmoniques, on propose d'appliquer l'algorithme de sélection JMI décrit dans la section III-2. Cet algorithme de type "Filters" basé sur le critère de maximisation de l'IM est largement utilisé pour la sélection de paramètres pertinents à

partir d'un ensemble initial de grande dimension. Dans le deuxième système IAED, basé sur l'analyse par ondelettes, on propose d'appliquer l'algorithme de sélection de type "Warppers" décrit dans la section III-3-2. Cet algorithme est appliqué généralement pour la sélection des paramètres pertinents à partir d'un ensemble initial de faible dimension.

IV-3 Expériences et résultats

Dans cette section, nous présentons les différentes expériences menées pour rechercher les bonnes configurations des deux types de systèmes IAED basés sur le classificateur KNN (en prenant la configuration par défaut k égal à 1 et la distance euclidienne). Ainsi, une partie de ce chapitre sera consacrée au système IAED basé sur l'analyse harmonique et une autre partie sera consacrée au système IAED basé sur l'analyse par ondelette. Également, une étude comparative entre les performances des deux systèmes sera présentée.

IV-3-1 Système IAED basé sur l'analyse harmonique

Dans cette section, nous présentons les différentes expériences effectuées pour évaluer les performances de notre système IAED basé sur l'extraction des paramètres statistiques. Plus particulièrement, nous allons étudier les points suivants:

- La meilleure combinaison des types de paramètres statistiques à savoir la moyenne, l'écart type, l'asymétrie et l'aplatissement;
- La recherche du nombre optimal de vecteurs de paramètres statistiques par seconde;
- L'effet de l'organisation et la distribution de la base de données, ainsi que l'utilisation d'autres critères d'évaluation des performances plus particulièrement la sensibilité, la précision, le F-score et la matrice de confusion;
- La réduction de la dimensionnalité basée sur l'application d'un algorithme de sélection de type "Filters".

IV-3-1-1 Combinaison optimale des paramètres statistiques

L'objectif de cette expérience est de chercher la meilleure combinaison, en terme de taux de classification, des quatre types de paramètres statistiques. En effet, chaque signal du courant électrique est converti en un seul vecteur de 1000 paramètres statistiques estimés sur toute la séquence de vecteurs de paramètres STFS dont chaque vecteur est composé de 250 paramètres STFS. Par conséquent, la classification KNN a été effectuée sans l'application de la règle de vote. Dans le but d'étudier l'effet de la dimensionnalité, nous allons également extraire des vecteurs statistiques à partir des séquences de vecteurs composé chacun de 50 paramètres harmoniques STFS.

Le tableau IV-3 montre les 15 combinaisons possibles des paramètres statistiques et leurs Taux de classification CR (%) avec la prise de 50 ou 250 harmoniques.

Tableau IV-3 Meilleure combinaison de paramètres statistiques

Combinaison des paramètres statistiques	Nombre de paramètres STFS		CR en %	
	50 Harmoniques	250 Harmoniques	50 Harmoniques	250 Harmoniques
μ	50	250	86.96	84.73
σ	50	250	84.35	83.79
Sk	50	250	75.79	72.62
Ku	50	250	67.59	67.78
μ, σ	100	500	93.10	92.36
μ, Sk	100	500	75.79	72.62
μ, Ku	100	500	67.59	67.78
σ, Sk	100	500	75.79	72.62
σ, Ku	100	500	67.59	67.78
Sk, Ku	100	500	68.52	68.34
μ, Sk, Ku	150	750	68.52	68.34
μ, σ, Ku	150	750	67.59	67.78
σ, Sk, Ku	150	750	68.52	69.34
μ, σ, Sk	150	750	75.79	72.62
μ, σ, Sk, Ku	200	1000	68.52	68.34

Comme le montre les résultats du tableau IV-3 la meilleure combinaison des types de paramètres statistiques est celle constituée de la moyenne et de l'écart type. Cette

combinaison permet d'obtenir les taux de classification les plus élevés avec un CR de 93,10% pour le cas de 50 harmoniques et un CR de 92,36% pour le cas de 250 harmoniques. Cette combinaison sera choisie dans les expériences suivantes.

Il est à noter que l'application des modèles HMM à 7 états (chacun d'eux étant associé à un modèle GMM de 3 gaussiennes sur des séquences de vecteurs de 250 paramètres STFS [40]) donne un taux de classification de 93,30%. Ici, le temps de calcul est long et est égal à environ 59 secondes (sans tenir compte du temps d'extraction de paramètres). En revanche, le classificateur KNN (avec les paramètres statistiques μ, σ) donne un taux de classification de 92,36 % avec un temps de calcul de 0,41 seconde (sans tenir compte le temps d'estimation des paramètres statistiques). Là, les expériences ont été réalisées en utilisant un ordinateur portable d'un processeur Intel Core i3 de 6 Go de mémoire RAM sous environnement MATLAB. Ce dernier résultat montre que l'approche proposée présente un CR comparable à celui de l'approche HMM mais avec un temps de calcul minimal ce qui montre l'avantage de l'utilisation de notre système vis-à-vis du problème de la complexité.

En outre, le tableau IV-3 montre également que dans la plupart des combinaisons, l'utilisation de 50 harmoniques donne de meilleurs résultats CR en comparaison avec le cas de l'utilisation de 250 harmoniques. Ceci montre aussi l'avantage de la réduction de dimensionnalité permettant d'améliorer les performances du système en termes de complexité et de précision, en évitant éventuellement le phénomène de la malédiction. Par conséquent, la combinaison optimale conduisant aux vecteurs de 500 paramètres statistiques peut nécessiter une réduction de la dimensionnalité ce que nous allons voir par l'application de l'algorithme de sélection JMI.

IV-3-1-2 Configuration optimale du classificateur KNN

Cette expérience a pour objectif de rechercher la bonne configuration du classificateur KNN en déterminant la valeur optimale du nombre de vecteurs les plus proche voisins k ($k=1:10$), ainsi que le type de la distance (ou mesure de similarité) en terme du taux de classification. Cette expérience est réalisée dans les mêmes conditions de l'expérience précédente en prenant la combinaison optimale de type de paramètres statistiques (μ, σ). Le tableau IV-4 présente les résultats de classification pour les différentes valeurs

de k ainsi que pour les différentes distances (Euclidienne, Corrélation, Cosinus, Cityblock).

Tableau IV-4: Configuration optimale de classificateur KNN

K \ Distance	Euclidienne	Corrélation	Cosinus	Cityblock
1	92.36	82.49	81.75	88.82
2	92.36	82.49	81.75	88.82
3	91.99	81.56	80.81	89.38
4	91.62	83.24	81.19	89.01
5	88.64	81.00	77.65	87.15
6	88.82	81.19	78.77	87.52
7	86.21	78.02	78.02	85.10
8	85.66	79.32	78.02	85.66
9	84.54	78.02	78.02	85.47
10	82.86	78.95	77.28	85.47

A partir de ce tableau, les résultats montrent que le meilleur taux CR égal à 92.36 % est obtenu en choisissant $k=1$ et la distance euclidienne. On choisira cette configuration dans les expériences suivantes.

IV-3-1-3 Nombre optimal de vecteurs statistiques de vote

Comme nous l'avons déjà montré, nos systèmes IAED proposés sont basés sur l'extraction des paramètres statistiques estimés, initialement dans les premières expériences, à partir de la partie entière du signal. Cependant, le processus d'identification des appareils électriques se déroule de manière permanente, ce qui exige de limiter la durée de l'estimation des paramètres statistiques en découpant le signal en segments dont chacun sera converti en un vecteur de paramètres statistiques.

L'objectif de cette expérience est de déterminer le nombre de segments ou vecteurs statistiques N_{vs} par seconde. La tâche IAED consiste à classifier le signal d'entrée par l'application de la stratégie de règle de vote sur la séquence d'indices de classes obtenue par la classification KNN à partir de la séquence des vecteurs statistiques représentant le signal. Le tableau IV-5 montre le CR (obtenu avec 250 harmoniques), les différents

nombres de vecteurs statistiques Nvs par seconde, ainsi que le temps en secondes écoulé pour déterminer chaque nombre Nvs.

Tableau IV-5: Nombre optimal de vecteurs statistiques (par seconde) obtenu par la méthode des règles de vote, en utilisant 250 harmoniques.

Le nombre de vecteurs statistiques par seconde	CR (%)	Temps écoulé [seconds]
1	92.36	0.47
2	94.41	0.80
3	94.78	0.98
4	94.22	1.20
6	94.22	1.68
8	94.97	2.03
10	94.78	2.46
12	94.97	2.83
15	94.97	3.39
20	94.78	4.46
30	94.59	6.50
40	94.78	9.04
60	94.78	11.79

A partir de ce tableau, il est clair que 8, 12 et 15 vecteurs statistiques donnent le meilleur CR égal à 94,97 %, en utilisant 250 harmoniques. Ces nombres correspondent respectivement aux nombres de vecteurs instantanés STFS 15, 10 et 8. De plus, les résultats montrent que le temps de calcul (même configuration matérielle utilisée dans la section IV-3-1-1), pour le cas de l'utilisation de 8 vecteurs statistiques, donne le temps d'exécution le plus court (2,03 s) comparé à celui de 12 vecteurs statistiques (2,83 s) et celui de 15 vecteurs statistiques (3,39 s).

A partir de ces résultats, nous pouvons conclure que le nombre optimal de vecteurs statistiques par seconde est égal à huit (8). Ce nombre sera choisi dans les expériences suivantes.

IV-3-1-4 Effet de la taille des bases de données d'apprentissage et de test sur les performances d'identification

Dans les premières expériences, les performances du premier système IAED proposé sont évaluées en utilisant une répartition de données de 50% pour la base d'apprentissage et 50% pour la base de test. Dans le but d'étudier la robustesse du système en vue du changement de la répartition des bases de données d'apprentissage

et de test, nous allons évaluer le taux de classification pour différentes répartitions. Le tableau IV-6 montre le taux de classification CR pour différentes répartitions de la base de données PLAID.

Tableau IV-6: Effet de la répartition de la base de données sur les résultats de CR

Apprentissage/test	50%/50%	60%/40%	80%/20%
CR (%)	94.97	97.20	97.20

Le tableau IV-6 montre que l'augmentation de la taille de la base de données d'apprentissage à plus 60% améliore les performances du système IAED. Cela peut être justifié par l'ajout de quantité d'informations permettant de bien discriminer les différentes classes d'appareils.

IV-3-1-5 Evaluation approfondie des performances

Comme nous l'avons déjà dévoilé, le taux de classification (CR) donne une vision globale sur les performances des systèmes IAED. Pour des évaluations plus approfondies, nous utilisons d'autres métriques fréquemment utilisées dans la littérature [52], [53], [172], [54], telles que la sensibilité S , la précision P , le F-score F_s et la matrice de confusion. Ces métriques peuvent être considérées comme des mesures locales permettant d'évaluer les performances pour chaque appareil électrique [12], [52], [53]. Le tableau IV-7 montre les résultats obtenus, pour les différentes métriques d'évaluation, de chaque type d'appareil électrique en considérant les différentes répartitions de la base de données PLAID.

Tableau IV-7: Sensibilité S , précision P et F-score F_s de chaque type d'appareil électrique pour différentes répartitions de la base de données

Apprentissage % / Test % (NBRApp / NBRTTest)		50/50 (537/537)			60/40 (645/429)			80/20 (859/215)		
N	Type d'appareil	S	P	F_s	S	P	F_s	S	P	F_s
1	Lampe fluorescente compacte	95.45	97.67	96.55	100	97.22	98.59	100	100	100
2	Aspirateur	50.00	63.63	56.00	60	81.81	69.23	62.50	83.33	71.42
3	Sèche-cheveux	97.56	96.38	96.96	100	93.93	96.87	100.00	93.93	96.87
4	Micro-onde	100	100	100	98.21	100	99.09	100	100	100
5	Climatiseur	87.87	100	93.54	96.15	100	98.03	92.30	100	96.00
6	PC Portable	97.67	95.45	96.55	97.10	100	98.52	100	100	100
7	Réfrigérateur	100	100	100	100	100	100	100	100	100
8	Ampoule à lumière incandescence	94.73	90.00	92.30	100	97.87	98.92	100	95.83	97.87
9	Ventilateur	94.82	94.82	94.82	100	97.87	98.92	100	95.83	97.87
10	Machine à laver	92.30	100	96.00	100	100	100	100	100	100

11	Chauffe-eau	94.11	76.19	84.21	85.71	85.71	85.71	71.42	83.33	76.92
	Globalement	91.32	92.19	91.54	94.28	95.85	94.90	93.29	95.66	94.27

A partir de ce tableau, les résultats obtenus montrent les performances détaillées de notre système, pour chaque appareil électrique. Par exemple, les micro-ondes et les réfrigérateurs ont été classés avec des valeurs de 100 % pour les trois métriques et pour toutes les répartitions de la base de données. Ceci démontre ainsi une excellente performance d'identification. Les autres appareils ont été classés également comme ayant de performances très acceptables. En effet, à l'exception de l'aspirateur qui présente des résultats faibles tombant à 56 % pour le F-score, presque toutes des trois métriques sont supérieures à 90 %. De plus, le tableau IV-7 montre qu'une augmentation de la taille de la base de données d'apprentissage améliore les performances d'identification en termes de F-score dans la plupart des cas.

Le tableau IV-8 présente la matrice de confusion obtenue avec 50 % de la base de données d'apprentissage et 50 % de la base de données de test. Chaque colonne j de la matrice de confusion indique le nombre d'échantillons d'appareil de la classe j correspondant à la ligne $i=j$, correctement identifiée. Cette matrice indique également le nombre d'appareils de la classe j incorrectement identifiés et confondus avec la classe d'appareil correspondant à la ligne $i \neq j$. La valeur en gras est le CR global. Les points représentent les valeurs zéro.

Tableau IV-8: Table de la matrice de confusion

	Classes réelles												
	N	1	2	3	4	5	6	7	8	9	10	11	P (%)
Classes identifiées	1	84	2	97.67
	2	.	7	3	.	.	1	63.64
	3	.	2	80	.	1	96.39
	4	.	.	.	70	100
	5	29	100
	6	4	84	95.45
	7	19	100
	8	.	2	.	.	1	.	.	54	3	.	.	90.00
	9	.	1	.	.	1	.	.	.	55	1	.	94.83
	10	12	.	100
	11	.	2	2	.	1	16	76.19
	S (%)	95.45	50.00	97.56	100	87.88	97.67	100	94.74	94.83	92.31	94.12	94.97

A partir de la matrice de confusion, présentée dans le tableau IV-8, on peut donner les remarques suivantes:

- La classe de l'aspirateur présente de faibles valeurs de la précision et de la sensibilité. Ce résultat peut être justifié par la limitation du nombre d'échantillons de cette classe. Plus particulièrement, la faible valeur de 50% de la sensibilité est expliquée par le nombre important de FN obtenu par la confusion de cette classe avec les classes 3, 8, 9 et 11. Également, la faible valeur de 63.64% de la précision expliquée par le nombre important de FP obtenu par la confusion des classes 8 et 11 avec la classe de l'aspirateur;
- La classe de climatiseur est confondue avec plusieurs classes d'appareils, conduisant à une sensibilité de 87.88%. Par contre, aucun appareil n'est confondu avec cette classe, conduisant à une précision de 100%(FP=0);
- La classe de chauffe-eau est identifiée avec une sensibilité acceptable de 94.12%. Cependant, la confusion de plusieurs classes d'appareils avec cette classe conduit à une précision de 76.19%;
- La plupart des appareils sont identifiés avec une sensibilité et une précision acceptables dépassent 90%.

IV-3-1-6 Sélection des paramètres statistiques STFS

Cette expérience consiste à sélectionner les paramètres pertinents à partir de l'ensemble de 500 paramètres statistiques STFS, pour la tâche d'identification des 11 classes d'appareils. Cette sélection est basée sur l'algorithme JMI. L'estimation de l'IM est basée sur l'approche d'histogramme. L'obtention des histogrammes se base sur la discrétisation des paramètres en utilisant la formule de Sturges pour le choix du nombre de "bins".

La figure IV-9 présente l'évolution du taux de classification pour les 500 paramètres statistiques sélectionnés par la stratégie de sélection JMI, en tenant compte des différentes répartitions de la base de données.

Les courbes de la figure IV-6 permettent de faire les remarques suivantes:

- La sélection des paramètres statistiques pertinents permet de réduire considérablement la dimension des vecteurs de paramètres avec un taux de classification CR supérieur ou égal à celui obtenu en considérant l'ensemble de tous les paramètres statistique (500 paramètres), et ce quel que soit la répartition de la base de données;

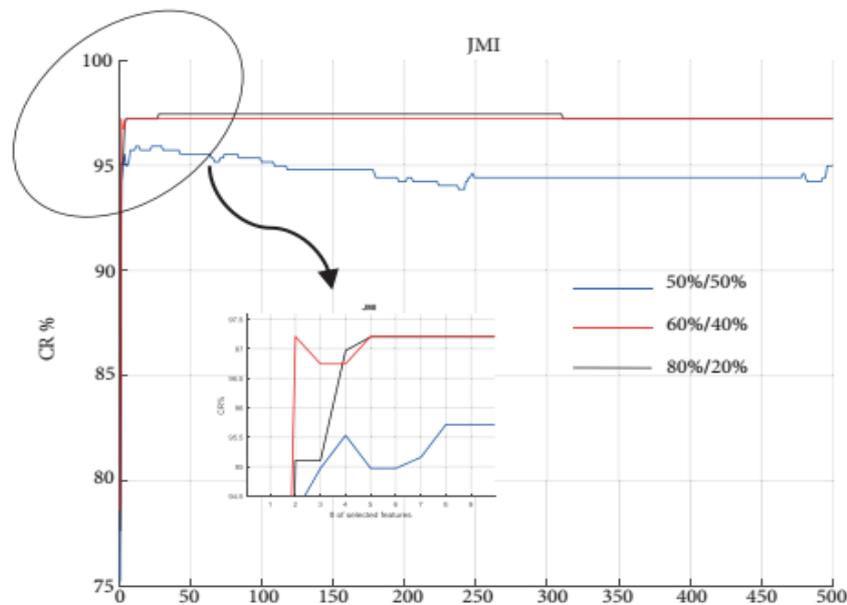


Figure IV-9: CR (%) pour la sélection de paramètres par la stratégie JMI avec les trois répartitions de la base de données.

- Dans le cas de la répartition 50%/50%, les résultats nous montrent que le CR augmente jusqu'à une valeur maximale de 95.90 % avec un nombre très réduit des paramètres (12 paramètres). Ensuite, la courbe subit une chute suivie d'un plateau à partir du nombre de paramètres sélectionnés qui est approximativement égal à 250. Ce comportement peut être éventuellement expliqué par le phénomène de la malédiction. L'absence de ce phénomène dans les autres courbes peut être justifiée par l'augmentation du nombre d'échantillons de la base d'apprentissage.

Le tableau IV-9 montre le taux de classification des 10 premiers paramètres statistiques sélectionnés et le taux de classification correspondant aux 500 paramètres avec les trois répartitions de la base de données. Le NBR indique le nombre de paramètres sélectionnés et le NBS indique le rang de paramètre sélectionné.

Les résultats de ce tableau indiquent que seulement 5 paramètres sont suffisants pour atteindre un CR supérieur ou égal à celui atteint en utilisant 500 paramètres et ce quel que soit la répartition de la base de données. Cela montre qu'une réduction de la dimensionnalité peut être obtenue sans aucune perte des performances.

Tableau IV-9: Taux de classification CR (%) des 10 premiers paramètres sélectionnés comparés aux 500 paramètres

NBR	50%/50%		60%/40%		80%/20%	
	NBS	CR(%)	NBS	CR (%)	NBS	CR (%)
1	1	75.23	1	77.62	1	78.60
2	3	94.22	5	95.10	3	97.20
3	5	94.97	7	95.10	7	96.74
4	7	95.53	3	96.96	5	96.74
5	9	94.97	9	97.20	9	97.20
6	11	94.97	11	97.20	11	97.20
7	13	95.15	15	97.20	15	97.20
8	15	95.71	13	97.20	13	97.20
9	39	95.71	17	97.20	17	97.20
10	43	95.71	19	97.20	19	97.20
500	252	94.71	377	97.20	367	97.20

Les résultats montrent également que les paramètres correspondant à la moyenne statistique des harmoniques de rang impair sont principalement les paramètres les plus pertinents pour cette tâche IAED. De plus, les résultats montrent que le choix de la répartition de la base de données n'a aucun effet sur les huit premiers paramètres sélectionnés. Malgré cela, elle présente un léger effet sur l'ordre de sélection.

IV-3-2 Système IAED basé sur l'analyse par ondelettes

Dans cette section, on présente des expériences dans lesquelles nous cherchons la configuration optimale de notre système IAED basé sur l'extraction de descripteurs DWE, LWE et WCC. La conception du système se base sur la répartition de la base de données PLAID [165]. en 50% pour la base d'apprentissage et 50% pour la base de test. Les détails de cette répartition sont illustrés sur le tableau IV-2. Dans ce qui suit nous allons discuter les points suivants:

- La comparaison des performances du système IAED pour les différents descripteurs, à savoir : le STFS, le LOG_E, le DWE, le LWE et le WCC;
- La recherche du nombre optimal de vecteurs statistiques par seconde;
- La recherche de la meilleure durée de la fenêtre d'analyse;
- La recherche de la meilleure combinaison de la fonction Ondelette mère avec son niveau de décomposition;

- L'évaluation approfondie des performances du système, en utilisant d'autres métriques, telles que la sensibilité, la précision, le f-score et la matrice de confusion;
- La sélection des paramètres pertinents des différents descripteurs en utilisant l'approche "Warppers".

IV-3-2-1 Comparaison des performances du système IAED

L'objectif de cette expérience est de comparer les performances du système IAED pour les descripteurs STFS, LoG_E, DWE, LWE et WCC. De plus, cette expérience nous permet de rechercher le nombre optimal de vecteurs statistiques par seconde pour chaque descripteur. Pour le calcul des descripteurs basés sur l'analyse par ondelettes, nous avons choisi les statistiques de la moyenne et de l'écart type. Ces derniers sont estimés sur des vecteurs de paramètres extraits à partir de l'analyse par ondelettes en prenant la fonction d'ondelette mère DB5 avec une décomposition jusqu'au le niveau 5 [40].

Le tableau IV-10 montre le taux de classification CR des différents descripteurs pour différentes valeurs du nombre de vecteurs statistiques par seconde.

Tableau IV-10: Comparaison des performances en terme de CR en % des différents types de descripteurs pour différentes valeurs du nombre de vecteurs statistiques.

NBR des vecteurs statistiques par seconde	LOG_E (2 paramètres)	STFS (500 paramètres)	DWE (12 paramètres)	LWE (12 paramètres)	WCC (12 paramètres)
1	82.12	92.36	73.55	92.92	92.92
2	84.73	94.41	80.81	93.29	92.55
3	83.98	94.78	81.37	93.66	93.85
4	84.17	94.22	86.77	93.29	92.92
6	84.73	94.22	87.15	92.92	93.10
8	83.98	94.97	88.26	94.04	94.04
10	85.66	94.78	87.89	94.04	93.85
12	85.47	94.97	87.89	94.78	94.59
15	85.47	94.97	87.70	94.78	94.97
20	85.28	94.78	88.26	94.04	94.41
30	84.54	94.59	89.19	94.41	94.04
40	83.79	94.78	88.64	94.59	93.48
60	83.05	94.78	88.45	93.85	93.48

A partir de ces résultats, nous pouvons donner les remarques suivantes :

- Le CR maximal est obtenu, pour la majorité des descripteurs, en choisissant 15 vecteurs statistiques par seconde. Ce choix donne respectivement 85,47%, 94,97%, 87,70%, 94,78% et 94,97% pour les descripteurs, LOG_E, STFS, DWE, LWE et WCC. On peut observer que le descripteur WCC, basé sur l'utilisation des vecteurs de 12 paramètres statistiques, donne le même CR du descripteur STFS utilisant des vecteurs de 500 paramètres statistiques. Ce résultat montre l'avantage, en termes de taux de réduction de la dimensionnalité, issu de l'utilisation de l'analyse par ondelettes;
- Le descripteur LOG_E, basé sur vecteurs de deux paramètres statistiques donne un résultat acceptable correspondant à 85,47%. Ce résultat montre la pertinence de ce descripteur qui permettra d'améliorer probablement le CR par sa concaténation avec les d'autres descripteurs.
- D'une manière générale, le descripteur WCC offre le meilleur compromis entre le CR et la réduction de la dimensionnalité et ce quel que soit le nombre des vecteurs statistiques par seconde.

IL a été noté que ces résultats peuvent être améliorés en étudiant le choix de la durée de la fenêtre d'analyse ainsi que le choix de la fonction mère et le niveau de la décomposition dyadique. Dans les expériences suivantes, nous allons chercher les meilleurs choix de ces paramètres de configuration du système.

IV-3-2-2 Durée optimale de la fenêtre d'analyse

Dans cette expérience, nous prenons la configuration optimale obtenue dans les expériences précédentes, à savoir la famille d'ondelettes DB5 avec un niveau de décomposition égal à 5, ainsi que le nombre de 15 vecteurs statistiques par seconde. Ici, nous varions la durée de la fenêtre d'analyse de 1 jusqu'à 12 fois la durée initiale (200 ms) [40] et nous évaluons ensuite le taux de classification pour chaque durée en considérant les descripteurs LOG_E, DWE, LWE et WCC. Le Tableau IV-11 présente le CR pour différents descripteurs et différentes durées de la fenêtre d'analyse.

Tableau IV-11 : Taux de classification CR pour différents descripteurs et pour différentes durées de la fenêtre d'analyse

Cycles	LOG_E	DWE	LWE	WCC
1	85.47	87.70	94.78	94.97
2	82.49	91.99	96.46	96.64
3	84.17	93.85	97.39	97.02
4	83.61	94.04	97.57	98.13
5	81.93	94.41	97.76	97.57
6	81.93	94.41	97.39	97.39
7	78.77	94.41	97.76	97.76
8	74.11	95.15	97.02	97.20
9	74.30	95.15	96.83	97.02
10	74.48	94.97	96.64	96.46
11	72.99	93.29	95.71	96.08
12	67.22	90.13	92.73	92.92

À partir du tableau IV-11, nous remarquons que le nombre de quatre cycles donne un CR maximal de 98,13% pour les paramètres WCC. Ce résultat montre que l'analyse par ondelettes fonctionne mieux avec des durées plus grandes de la durée initiale de la fenêtre d'analyse. En contrepartie, on remarque que le choix d'une durée plus grande diminue les performances du système basé sur le descripteur LOG_E. Le meilleur choix de la durée de la fenêtre d'analyse sera pris en considération pour les expériences suivantes.

IV-3-2-3 Choix de l'ondelette mère et du niveau de décomposition dyadique

L'objectif de cette expérience est de chercher la meilleure fonction mère parmi les trois familles de fonctions mères couramment utilisées, à savoir Daubechies, Coiflet et Symlet. De plus, nous allons chercher pour chaque fonction mère le niveau de décomposition dyadique optimale. Les tableaux IV-12, IV-13 et IV-14 présentent les CR maximaux obtenus et les niveaux de décomposition correspondants en considérant plusieurs ordres de chaque famille de la fonction mère.

A partir des résultats obtenus dans les tableaux IV-12, IV-13 et IV-14, on peut donner les remarques suivantes :

- La fonction d'ondelette mère DB5 avec le niveau de décomposition 5 donne un meilleur CR égal à 98,13 %. Par contre, les fonctions mères Coif1 avec le niveau de décomposition 4 et Sym4 avec le niveau de décomposition 4 donnent le même CR

qui est égal à 97,95 %. Ce résultat confirme le bon choix de la fonction mère DB5 avec son niveau de décomposition, utilisé dans la référence [40].

Tableau IV-12: CR% maximal obtenu pour chaque ordre N de la famille Daubechies et son niveau de décomposition correspondant

DB(N)	niveau (P)	CR%
DB1	5	94.97
DB2	5	97.57
DB3	4	97.95
DB4	5	97.39
DB5	5	98.13
DB6	5	97.20
DB7	7	97.20
DB8	5	96.83

Tableau IV-13 : CR% maximal obtenu pour chaque ordre N de la famille Coiflet et son niveau de décomposition correspondant

COIFN	Coif1	Coif2	Coif3	Coif4	Coif5
Niveau (P)	4	4	5/7	7	7
CR%	97.95	97.39	97.20	97.20	96.27

Tableau IV-14 : CR% maximal obtenu pour chaque ordre N de la famille Symlet et son niveau de décomposition correspondant

SYMN	Level (P)	CR%
Sym1	5	94.97
Sym2	5	97.57
Sym3	4	97.95
Sym4	4	97.76
Sym5	6	97.20
Sym6	6	97.39
Sym7	5	97.57
Sym8	5	97.20

IV-3-2-4 Évaluation approfondie des performances

Dans cette expérience, nous appliquons les différentes métriques (sensibilité S, précision P et F-score F) pour une évaluation approfondie des performances. Les valeurs de ces métriques, pour chaque type d'appareil électrique, sont rapportées dans le tableau IV-15. Les résultats sont obtenus en prenant 15 vecteurs statistiques par seconde des descripteurs WCC avec des fenêtres d'analyse de durée 200 ms (quatre cycles) en utilisant la famille d'ondelettes DB avec le niveau de décomposition 5.

Tableau IV-15: Sensibilité S, précision P et F-score Fs de chaque type d'appareil électrique

NBR	Type d'appareil	Sensitivité (%)	Précision (%)	F-score (%)
1	Lampe fluorescente compacte	100	96.70	98.32
2	Aspirateur	78.57	84.62	81.48
3	Sèche-cheveux	100	98.80	99.39
4	Micro-onde	100	100	100
5	Climatiseur	87.87	100	93.54
6	PC Portable	97.67	100	98.82
7	Réfrigérateur	100	95	97.43
8	Ampoule à lumière incandescence	100	95	97.43
9	Ventilateur	100	100	100
10	Machine à laver	100	100	100
11	Chauffe-eau	94.11	100	96.96
Globalement		96.20	97.28	96.67

Les résultats obtenus à partir du tableau IV-15 montrent les performances détaillées de notre système d'identification, pour chaque appareil électrique. Par exemple, le micro-ondes, la machine à laver et le ventilateur ont démontré une excellente performance d'identification. En effet, ces trois appareils ont été classés avec des valeurs de 100 % pour les trois mesures. Pour les autres appareils, les performances sont acceptables et les mesures sont presque toutes supérieures à 90 %, à l'exception de l'aspirateur qui présente un résultat faible mais amélioré par rapport aux résultats obtenus en utilisant le descripteur STFS.

Le tableau IV-16 montre la matrice de confusion pour une configuration optimale.

Tableau IV-16 Matrice de confusion obtenue avec la configuration optimale

N	1	2	3	4	5	6	7	8	9	10	11	P (%)
1	88	.	.	.	1	2	96.70
2	.	11	.	.	1	1	84.62
3	.	.	82	.	1	98.80
4	.	.	.	70	100
5	29	100
6	84	100
7	1	.	19	95
8	.	3	57	.	.	.	95
9	58	.	.	100
10	13	.	100
11	16	100
S (%)	100	78.57	100	100	87.88	97.67	100	100	100	100	94.11	98.13

A partir des tableaux IV-16 et IV-8, on peut remarquer que l'utilisation du descripteur WCC avec la dernière configuration améliore les performances du système en termes des trois métriques pour la majorité des classes d'appareils avec une diminution considérable du FN et du FP.

IV-3-4 Résultats de la sélection des paramètres

Cette partie est consacrée à la sélection des paramètres statistiques les plus pertinents à partir d'un ensemble initial de 12 paramètres statistiques extraits en utilisant la dernière configuration du système. Nous rappelons ici que la configuration optimale est basée sur l'utilisation de la fonction d'ondelette mère DB5 avec un niveau de décomposition égal à 5. Là, la sélection, basée sur l'approche "Wrappers", est appliquée pour chaque descripteur (DWE, LWE et WCC). Pour le descripteur DWE, le sous-ensemble de paramètres initiaux F_D est donné comme suit :

$$F_D = \{M_{ap}, M_{d5}, M_{d4}, M_{d3}, M_{d2}, M_{d1}, Std_{ap}, Std_{d5}, Std_{d4}, Std_{d3}, Std_{d2}, Std_{d1}\}.$$

où M_{ap} et Std_{ap} représentent respectivement la moyenne et l'écart type des coefficients d'approximation DWT et M_{di} et Std_{di} représentent respectivement la moyenne et l'écart type des coefficients de détails DWT au niveau i .

L'ordre initial des paramètres est pris de basses fréquences aux hautes fréquences.

Pour le descripteur LWE, Le sous-ensemble initial des paramètres F_L est donné comme suit :

$$F_L = \{LM_{ap}, LM_{d5}, LM_{d4}, LM_{d3}, LM_{d2}, LM_{d1}, LStd_{ap}, LStd_{d5}, LStd_{d4}, LStd_{d3}, LStd_{d2}, LStd_{d1}\}, \text{ Où le préfixe } L \text{ indique le descripteur LWE.}$$

Pour le descripteur WCC, le sous-ensemble initial des paramètres F_W est donné comme suit : $F_W = \{wcc_1, wcc_2, wcc_3, wcc_4, wcc_5, wcc_6, wcc_7, wcc_8, wcc_9, wcc_{10}, wcc_{11}, wcc_{12}\}$.

Le tableau IV-17 donne le CR de l'ensemble composé de l'ensemble de paramètres déjà sélectionnés avec le paramètre sélectionné à l'itération j .

A partir du tableau IV-17, nous pouvons faire les remarques suivantes :

- La sélection des paramètres améliore le CR qui atteint une valeur de 98,13 %, en utilisant uniquement six paramètres LWE. De plus, il dépasse le CR obtenu avec l'ensemble initial des paramètres (12 paramètres) ;

Tableau IV-17: CR% en fonction des paramètres sélectionnés.

Sel est le numéro du paramètre sélectionné et Feat est le nom du paramètre sélectionné à l'itération j.

J	DWE (12 paramètres)			LWE (12 paramètres)			WCC (12 paramètres)		
	Sel	Feat	CR%	Sel	Feat	CR%	Sel	Feat	CR%
1	5	M _{d2}	64.80	5	LM _{d2}	65.17	1	wcc ₁	62.19
2	6	M _{d1}	90.68	3	LM _{d4}	91.99	4	wcc ₄	91.62
3	3	M _{d4}	94.78	4	LM _{d3}	96.27	3	wcc ₃	96.27
4	4	M _{d3}	96.08	2	LM _{d5}	97.02	2	wcc ₂	97.39
5	9	Std _{d4}	96.83	6	LM _{d1}	97.95	5	wcc ₅	97.76
6	11	Std _{d2}	97.02	8	LStd _{d5}	98.13	6	wcc ₆	97.95
7	10	Std _{d3}	96.83	7	LStd _{ap}	98.13	8	wcc ₈	97.95
8	12	Std _{d1}	96.83	1	LM _{ap}	97.95	11	wcc ₁₁	97.95
9	7	Std _{ap}	96.46	11	LStd _{d2}	97.95	9	wcc ₉	97.95
10	8	Std _{d5}	96.83	12	LStd _{d1}	97.95	7	wcc ₇	98.13
11	2	M _{d5}	95.15	10	LStd _{d3}	97.76	10	wcc ₁₀	98.13
12	1	M _{ap}	94.04	9	LStd _{d4}	97.57	12	wcc ₁₂	98.13

- La sélection des paramètres du descripteur WCC permet d'obtenir le même CR de l'ensemble initial en utilisant 10 paramètres ;
- La sélection des paramètres du descripteur DWE améliore considérablement le CR qui atteint une valeur de 97,02 % en utilisant uniquement six paramètres. En outre, il dépasse le CR de 94,04 % obtenu en utilisant l'ensemble initial des paramètres (12 paramètres);
- Les résultats montrent la pertinence des paramètres calculés à partir de la moyenne statistique des descripteurs DWE et LWE.

Nous pouvons conclure, à partir de cette étude, que le descripteur LWE avec seulement six paramètres, donne le plus petit sous-ensemble sélectionné expliquant mieux les différentes classes d'appareils électriques considérés dans la base de données PLAID.

Dans la section (IV-3-2-1), les résultats ont montré la pertinence du descripteur LOG_E. Ainsi, la concaténation de ce descripteur avec chaque descripteur (DWE, LWE ou WCC) et la sélection des paramètres pertinents à partir de l'ensemble de 14 paramètres obtenu par cette concaténation peut probablement améliorer le CR. Les deux

paramètres M_E et Std_E ajoutés du descripteur LOG_E représentent respectivement la moyenne et l'écart du logarithme d'énergie.

Le tableau IV-18 donne le CR et le paramètre sélectionné à chaque itération j.

Tableau IV-18 CR% en fonction des paramètres sélectionnés avec concaténation de chaque descripteur (DWE, LWE ou WCC) avec le descripteur LOG_E.

J	DWE_LOG_E (14 features)			LWE_LOG_E (14 features)			WCC_LOG_E (14 features)		
	Sel	Feat	CR%	Sel	Feat	CR%	Sel	Feat	CR%
1	13	M_E	70.01	13	M_E	70.01	13	M_E	70.01
2	1	M_{ap}	86.21	5	LM_{d2}	94.04	2	wcc2	93.85
3	14	Std_E	89.19	2	LM_{d5}	97.39	1	wcc1	97.95
4	2	M_{d5}	90.87	4	LM_{d3}	97.76	3	wcc3	98.32
5	3	M_{d4}	91.06	3	LM_{d4}	98.13	10	wcc10	98.51
6	8	Std_{d5}	91.24	1	LM_{ap}	98.13	5	wcc5	98.51
7	4	M_{d3}	91.24	7	$LStd_{ap}$	98.13	6	wcc6	98.51
8	5	M_{d2}	91.24	9	$LStd_{d4}$	98.13	11	wcc11	98.51
9	6	M_{d1}	91.24	6	LM_{d1}	97.95	7	wcc7	98.32
10	9	Std_{d4}	91.24	8	$LStd_{d5}$	97.76	9	wcc9	98.32
11	10	Std_{d3}	91.24	10	$LStd_{d3}$	97.76	12	wcc12	98.32
12	12	Std_{d1}	91.24	11	$LStd_{d2}$	97.76	8	wcc8	98.13
13	7	Std_{ap}	91.04	12	$LStd_{d1}$	97.76	14	Std_E	98.13
14	11	Std_{d2}	91.06	14	Std_E	97.76	4	wcc4	97.76

A partir du tableau IV-18, on peut faire les remarques suivantes :

- Le premier paramètre sélectionné, pour chaque concaténation, est le paramètre M_E du descripteur LOG_E, qui contribue considérablement à expliquer les classes d'appareils ;
- La sélection du paramètre M_E du descripteur LOG_E avec quatre paramètres du descripteur WCC permet d'obtenir un CR maximal de 98,51%. Ce résultat montre l'avantage de la concaténation et la sélection des paramètres dans l'amélioration du CR et la réduction de la dimensionnalité;
- Dans le cas de la concaténation des descripteurs LOG_E et WCC, la sélection de trois paramètres conduit à un CR de 97.95% dépassant par conséquent le CR obtenu de l'ensemble initial comportant 14 paramètres.

On peut conclure que la concaténation du descripteur LOG_E avec les descripteurs basés sur l'analyse par ondelettes, permet d'améliorer les performances, en termes de précision et complexité, du système IAED.

IV-4 Conclusion

Dans ce chapitre, nous avons présenté les deux systèmes d'identification des appareils électriques IAED proposés ainsi que les différentes expériences menées pour chercher la configuration optimale pour chaque système. Ce chapitre est divisé en deux parties.

Comme nous l'avons vu tout au long de ce chapitre les deux systèmes sont basés sur l'utilisation des paramètres statistiques ainsi que la combinaison du classificateur KNN avec la stratégie de règle de vote. Cependant, ces systèmes se différencient dans l'étape d'extraction des paramètres. En effet, le premier système est basé sur l'extraction des paramètres statistiques du descripteur STFS et le deuxième est basé sur l'extraction des paramètres statistiques des descripteurs issus de l'analyse par ondelettes.

Plusieurs expériences ont été menées afin de rechercher la bonne configuration de chaque système permettant d'améliorer les performances avec dimensionnalité réduite.

Les résultats obtenus, pour le premier système proposé, montrent que les paramètres statistiques, basés sur la moyenne et l'écart type avec 8 vecteurs statistiques par seconde, donne de meilleures performances. De plus, les résultats de la sélection des paramètres nous ont montré que le sous-ensemble, comportant les 5 premiers paramètres sélectionnés « STFS » d'ordre impair, est suffisant pour bien expliquer les 11 classes d'appareils.

Concernant le deuxième système proposé, les résultats obtenus ont montré la pertinence des paramètres statistiques WCC extraits à partir de l'analyse par ondelettes utilisant la fonction mère DB5 avec une décomposition jusqu'au niveau 5. De plus, la concaténation des paramètres du descripteur LOG_E avec les paramètres du descripteur WCC améliore les performances du système.

Globalement, les résultats obtenus ont montré l'amélioration des performances en termes de taux de classification et de complexité de la méthode d'extraction des paramètres basée sur l'analyse par ondelettes.

CONCLUSION GENERALE

Conclusion Générale

Comme nous l'avons vu tout au long de cette thèse, l'identification des appareils électriques domestiques est une fonction primordiale intégrée dans les compteurs électriques intelligents. Ces compteurs intègrent généralement dans leur fonctionnement une fonction importante qui consiste à fournir des informations en temps réel sur l'identité de l'appareil électrique et l'intervalle de son fonctionnement ainsi que l'énergie consommée. Nous avons montré, pendant la réalisation de ce travail de Doctorat, que cette tâche peut être effectuée en concevant des systèmes d'identification des appareils électriques à partir de l'exploitation des signaux électriques. L'objectif alors était d'identifier la classe d'appareil en état de marche à partir des mesures faites au niveau du compteur d'énergie électrique principal.

Dans ce contexte, un nouveau système pour l'identification d'appareils électriques est proposé. Nos premières idées clés ont été inspirées de l'utilisation des paramètres statistiques des harmoniques et l'application du classificateur k-NN combiné avec la méthode des règles de vote. A cet effet, l'extraction des paramètres des signaux de courant a été effectuée sur les statistiques des coefficients STFS notamment la moyenne, l'écart type, l'asymétrie et l'aplatissement. Les premiers résultats obtenus, évalués sur la base de données PLAID, ont montré que la combinaison des paramètres « moyennes » et « écarts types » fourni des performances optimales en termes de taux de classification CR. Ces résultats ont été améliorés, avec un gain de 2,6%, par l'application de la règle de vote avec le choix du nombre optimal de vecteurs de vote.

Afin de réduire la haute dimensionnalité (500 paramètres), nous avons également appliqué un algorithme de sélection des paramètres basé sur la stratégie d'information mutuelle JMI. L'application de cette dernière a montré que la procédure de sélection des paramètres est devenue efficace pour réduire le nombre de paramètres tout en améliorant les valeurs du taux CR.

Nos résultats ont également montré la robustesse, en termes de plusieurs critères d'évaluation des performances, du système proposé par rapport à différentes distributions d'apprentissage et de test de la base de données. En effet, nous avons démontré que la configuration optimale de ce système augmente les valeurs de CR jusqu'à 94,97 %. Cette valeur a été obtenue en sélectionnant les trois premiers paramètres, ce qui explique l'efficacité de l'algorithme de sélection JMI utilisé.

Nous avons prouvé pareillement que la réduction de la dimensionnalité peut être obtenue différemment. Dans cette optique, nous avons proposé d'appliquer la méthode d'extraction des paramètres statistiques estimés à partir des descripteurs basés sur l'analyse par ondelettes. Nous avons aussi appliqué la sélection des paramètres basée sur l'approche "*Wrappers*" pour étudier la pertinence de ces paramètres ainsi que les paramètres du descripteur LOG_E. Les résultats obtenus ont montré l'amélioration du taux de classification qui a atteint une valeur maximale de 98,51% en utilisant une configuration optimale avec seulement 5 paramètres.

Finalement, nous pouvons envisager les perspectives suivantes :

- L'utilisation d'autres types de classificateurs tels que les réseaux de neurones artificiels (ANN), l'analyse discriminante linéaire (LDA), la machine à vecteurs de support (SVM)...etc.
- L'évaluation des performances du système d'identification sur une base de données de taille plus large.

Bibliographie

- [1] Nait-Meziane M, Hacine-Gharbi A, Ravier P, Lamarque G, Le Bunetel JC et al. HMM-based transient and steady state current signals modeling for electrical appliances identification. In: 5th International Conference on Pattern Recognition Applications and Methods; Rome, Italy; 2016. pp. 670-677.
- [2] Patel SN, Robertson T, Kientz JA, Reynolds MS, Abowd GD. At the flick of a switch: detecting and classifying unique electrical events on the residential power line. In: International Conference on Ubiquitous Computing; Innsbruck, Austria; 2007. pp. 271–288.
- [3] H.-H. Chang, C.-L. Lin, and J.-K. Lee. Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms. in Computer Supported Cooperative Work in Design (CSCWD), 2010 14th International Conference on. IEEE, 2010, pp. 27–32.
- [4] Nait-Meziane M, Hacine-Gharbi A, Ravier P, Lamarque G, Le Bunetel JC et al. Electrical appliances identification and clustering using novel turn-on transient features. In: 6th International Conference on Pattern Recognition Applications and Methods; Porto, Portugal; 2017. pp. 647-654.
- [5] Ghazali, F., Hacine-Gharbi, A., Ravier, P. and Mohamadi, T. ‘Extraction and selection of statistical harmonics features for electrical appliances identification using KNN classifier, combined with voting rules method’, Turkish Journal of Electrical Engineering & Computer Sciences, 2019, Vol. 27, No. 4, pp.2980–2997.
- [6] Ghazali, F., Hacine-Gharbi, A. and Ravier, P. ‘Statistical features extraction based on the discrete wavelet transform for electrical appliances identification’, presented at the International Conference of Intelligent Systems and Pattern Recognition (ISPR2020), Hammamet, Tunisia, 6–18 October, 2020.
- [7] Ghazali, F., Hacine-Gharbi, A. and Ravier, P. ‘Selection of statistical wavelet features using a wrapper approach for electrical appliances identification based on a KNN classifier combined with voting rules method’, Int. J. Computational Systems Engineering, 2021, Vol. 6, No. 5, pp.220–230

- [8]** W. Wichakool, Z. Remscrim, U.A. Orji, S.B. Leeb, Smart metering of variable power loads, *IEEE Trans. Smart Grid* 6 (1) (2015) 189–198.
- [9]** I. Abubakar, S. Khalid, M. Mustafa, H. Shareef, M. Mustapha, Application of load monitoring in appliances' energy management — a review, *Renew. Sustain. Energy Rev.* 67 (2017) 235–245.
- [10]** S.S. Hosseini, K. Agbossou, S. Kelouwani, A. Cardenas, Non-intrusive load monitoring through home energy management systems: a comprehensive review, *Renew. Sustain. Energy Rev.* 79 (2017) 1266–1274.
- [11]** Meehan, Paula. An efficient scalable time-frequency method for tracking energy usage of domestic appliances using a two-step classification algorithm. PhD thesis, Dublin City University. 2015.
- [12]** M. Zeifman and K. Roth, “Nonintrusive appliance load monitoring: Review and outlook,” *Consumer Electronics, IEEE Transactions on*, vol. 57, no. 1, pp. 76–84, 2011.
- [13]** Y. Du, L. Du, B. Lu, R. Harley, and T. Habetler, “A review of identification and monitoring methods for electric loads in commercial and residential buildings,” in *Energy Conversion Congress and Exposition (ECCE), IEEE*, 2010, pp. 4527–4533.
- [14]** H. Farhangi, “The path of the smart grid,” *Power and Energy Magazine, IEEE*, vol. 8, no. 1, pp. 18–28, 2010
- [15]** S. Nandi, H. Toliyat, and X. Li, “Condition monitoring and fault diagnosis of electrical motors-a review,” *Energy Conversion, IEEE Transactions on*, vol. 20, no. 4, pp. 719–729, 2005
- [16]** N. Mehala, “Condition monitoring and fault diagnosis of induction motor using motor current signature analysis,” , PHD thesis , national institute of technology Kurukshetra, India, 2010.

- [17] L. Norford and S. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy and Buildings*, vol. 24, no. 1, pp. 51–64, 1996.
- [18] W. C. Greene, "Evaluation of non-intrusive monitoring for condition based maintenance applications on US Navy propulsion plants", Master of Science in Naval Architecture and Marine Engineering and Master of Science in Mechanical Engineering at the Massachusetts Institute of Technology , University of Florida, 2005.
- [19] Vladimír Haasz, Kurosh Madani, " Advanced Data Acquisition and Intelligent Data Processing", River Publischers ,Aalborg, Denmark, 2014.
- [20] GW, Hart, Residential energy monitoring and computerized surveillance via utility power flows. *IEEE Technology and Society Magazine* , 8 (2)pp 12-16. 1989.
- [21] G. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
- [22] Sultanem F. Procédé et appareil d'analyse de signaux de courant et de tension en vue du repérage de charges à usage domestique. Brevet Français FR 2645968, 1990 (in French).
- [23] Bons M. Modèles à sources markovienne cachée appliqués à l'analyse non-intrusive des principaux usages domestiques de l'électricité. PhD, Université de Rennes I, Rennes, France, 1996 (in French).
- [24] Onada T, Nakano Y, Yoshimoto K. System and method for estimating power consumption of electric apparatus, and abnormality alarm system utilizing the same. US patent 6,816,078 B2, 2004
- [25] Leeb S. A conjoint pattern recognition approach to nonintrusive load monitoring. PhD, Massachusetts Institute of Technology, Cambridge, MA, USA, 1993.

- [26] Cole AI, Albicki A. Data extraction for effective nonintrusive identification of residential power loads. In: IEEE Instrumentation and Measurement Technology Conference; St. Paul, MN, USA; 1998. pp. 812-815.
- [27] Drenker S, Kader A. Nonintrusive monitoring of electric loads. IEEE Computer Applications in Power 1999; 12 (4): 47-51.
- [28] Chan WL, So ATP, Lai LL. Harmonics load signature recognition by wavelets transforms. In: International Conference on Electric Utility Deregulation and Restructuring and Power Technologies; London, UK; 2000. pp. 666-671.
- [29] Baranski M, Voss J. Nonintrusive appliance load monitoring based on an optical sensor. In: IEEE Bologna Power Tech Conference; Bologna, Italy; 2003. p. 8.
- [30] Chang HH, Lin CL, Yang HT. Load recognition for deferent loads with the same real power and reactive power in a nonintrusive load monitoring system. In: 12th IEEE International Conference on Computer Supported Cooperative Work in Design; Xi'an, China; 2008. pp. 1122-1127.
- [31] Du Y, Du L, Lu B, Harley R, Habetler T. A review of identification and monitoring methods for electric loads in commercial and residential buildings. In: IEEE Energy Conversion Congress and Exposition; Atlanta, GA, USA; 2010. pp. 4527-4533.
- [32] Carrie Armel K, Gupta A, Shrimali G, Albert A. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 2013; 52: 213-234.
- [33] Najmeddine H, El Khamlichi Drissi K, Pasquier C, Faure C, Kerroum K et al. States of art on load monitoring methods. In: 2nd IEEE International Conference on Power and Energy; Johor Baharu, Malaysia; 2008. pp. 1256- 1258.
- [34] Ridi A, Hennebert J. Hidden Markov models for ILM appliance identification. Procedia Computer Science 2014; 32: 1010-1015.
- [35] Mittelsdorf M, Hüwel A, Klingenberg T, Sonnenschein M. Submeter. based training of multi-class support vector machines for appliance recognition in home

- electricity consumption data. In: Proceedings of Smart Greens; Aachen, Germany; 2013. pp. 151-158.
- [36] M. Berges, E. Goldman, H. Matthews, and L. Soibelman, "Enhancing electricity audits in residential buildings with nonintrusive load monitoring," *Journal of Industrial Ecology*, vol. 14, no. 5, pp. 844–858, 2010.
- [37] H. H. Chang, K. L. Chen, Y. P. Tsai, W. J. Lee, A new measurement method for power signatures of nonintrusive demand monitoring and load identification, *IEEE Transactions on Industry Applications* 48 (2) (2012) 764–771. doi:10.1109/TIA.2011.2180497.
- [38] D. Srinivasan, W. S. Ng, A. C. Liew, Neural-network-based signature recognition for harmonic source identification, *IEEE Transactions on Power Delivery* 21 (1) (2006) 398–405. doi:10.1109/TPWRD.2005.852370.
- [39] Dat Nguyen T, Dong Do T, Ha Le M, Le NT, Benjapolakul W. Appliance classification method based on k-nearest neighbors for home energy management system. In: First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics; Bangkok, Thailand; 2019. pp. 53-56.
- [40] Hacine-Gharbi A, Ravier P. Wavelet cepstral coefficients for electrical appliances identification using hidden Markov models. In: 7th International Conference on Pattern Recognition Applications and Methods; Funchal, Portugal; 2018. pp. 541-549.
- [41] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, S. Y. Philip et al., "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.
- [42] T. Onoda, G. Rätsch et K.R. Müller. Applying Support Vector Machines and Boosting to a Non-Intrusive Monitoring System for Household Electric Appliances with Inverters, 2000.

- [43] S. Gupta, M. Reynolds, and S. Patel, "Electrisense: single-point sensing using EMI for electrical event detection and classification in the home," in Proceedings of the international conference on Ubiquitous computing. ACM, 2010, pp. 139–148
- [44] V. B. Surya Prasatha, Haneen Arafat Abu Alfeilate, Ahmad B. A. Hassanate, Omar Lasassmehe, Ahmad S. Tarawnehf, Mahmoud Bashir Alhasanath, Hamzeh S. Eyal Salmane. Effects of Distance Measure Choice on KNN Classifier Performance. Number 00, Mary Ann Liebert, : Big Data , 2019.
- [45] Hacine-Gharbi, A. "Sélection de paramètres acoustiques pertinents pour la reconnaissance de la parole". Phd thesis, Orléans University, France , 2012.
- [46] A. Cole and A. Albicki, "Nonintrusive identification of electrical loads in a three-phase environment based on harmonic content," in Instrumentation and Measurement Technology Conference, IMTC. Proceedings of the IEEE, vol. 1. IEEE, 2000, pp. 24–29.
- [47] J. Kolter and M. Johnson, "Redd: A public data set for energy disaggregation research," in Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, 2011.
- [48] Y. Wang, X. Hao, L. Song, C. Wu, Y. Wang, C. Hu, and L. Yu, "Tracking states of massive electrical appliances by lightweight metering and sequence decoding," in Proceedings of the Sixth International Workshop on Knowledge Discovery from Sensor Data. ACM, 2012, pp. 34–42.
- [49] A. Zoha, A. Gluhak, M. Nati, and M. A. Imran, "Low-power appliance monitoring using factorial hidden markov models", *IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing* pp. 527–532, 2013. doi: 10.1109/ISSNIP.2013.6529845
- [50] L. R. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition," Proceedings of the IEEE, vol. 77, no. 2, pp. 257–286, 1989.
- [51] S. R. Eddy, "What is a Hidden Markov Model?" Nat Biotech, vol. 22, no. 10, 2004.

- [52] De Paiva Penha D, Castro ARG. Home appliance identification for NILM systems based on deep neural networks. *International Journal of Artificial Intelligence and Applications* 2018; 9 (2): 69-80.
- [53] Kelly J, Knottenbelt W. Neural NILM: Deep neural networks applied to energy disaggregation. In: *2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*; Seoul, South Korea; 2015. pp. 55-64.
- [54] Djordjević S, Simić M. Nonintrusive identification of residential appliances using harmonic analysis. *Turkish Journal of Electrical Engineering & Computer Sciences* 2018; 26 (2): 780-791.
- [55] R. Kohavi et F. Provost. Glossary of Terms. Editorial for the Special Issue on Applications of Machine Learning and the Knowledge Discovery Process, vol. 30, pages 271-274, February 1998.
- [56] Thomas Bier . " Disaggregation of Electrical Appliances using Non-Intrusive Load Monitoring", thèse à Université de Haute-Alsace, Mulhouse École Doctorale Jean-Henri Lambert Laboratoire MIPS, 2014 télécharger depuis <https://tel.archives-ouvertes.fr/tel-01314432>
- [57] van Erp M, Vuurpijl L, Schomaker L. An overview and comparison of voting methods for pattern recognition. In: *Proceedings of Eighth International Workshop on Frontiers in Handwriting Recognition*; Niagara on the Lake, Canada; 2002. pp. 195-200.
- [58] Bilski P, Winiecki W. Non-intrusive appliance load identification with the ensemble of classifiers. In: *Proceedings of the International Workshop on Non-Intrusive Load Monitoring*; Vancouver, Canada; 2016. pp. 1-5.
- [59] A. Zoha, A. Gluhak, M. A. Imran, S. Rajasegarar, Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey, *Sensors* 12 (12) (2012) 16838–16866. doi:10.3390/s121216838.

- [60] J. T. Powers, B. Margossian, B. A. Smith, Using a rule-based algorithm to disaggregate end-use load profiles from premise-level data, *IEEE Computer Applications in Power* 4 (2) (1991) 42–47. doi:10.1109/67.75875.
- [61] L. Farinaccio, R. Zmeureanu, Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses, *Energy and Buildings* 30 (3) (1999) 245 – 259. doi:10.1016/S0378-7788(99)00007-9.
- [62] M. Marceau, R. Zmeureanu, Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings, *Energy Conversion and Management* 41 (13) (2000) 1389 – 1403. doi:10.1016/S0196-8904(99)00173-9.
- [63] F. Sultanem, Using appliance signatures for monitoring residential loads at meter panel level, *IEEE Transactions on Power Delivery* 6 (4) (1991) 1380–1385. doi:10.1109/61.97667.
- [64] A. G. Ruzzelli, C. Nicolas, A. Schoofs, G. M. P. O'Hare, Real-time recognition and profiling of appliances through a single electricity sensor, in: 2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2010, pp. 1–9. doi:10.1109/SECON.2010.5508244.
- [65] M. B. Figueiredo, A. de Almeida, B. Ribeiro, An Experimental Study on Electrical Signature Identification of Non-Intrusive Load Monitoring (NILM) Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 31–40. doi:10.1007/978-3-642-20267-4_4.
- [66] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, P. Armstrong, Power signature analysis, *IEEE Power and Energy Magazine* 1 (2) (2003) 56–63. doi:10.1109/MPAE.2003.1192027.

- [67] S. B. Leeb, S. R. Shaw, J. L. Kirtley, Transient event detection in spectral envelope estimates for nonintrusive load monitoring, *IEEE Transactions on Power Delivery* 10 (3) (1995) 1200–1210. doi:10.1109/61.400897.
- [68] C. Duarte, P. Delmar, K. W. Goossen, K. Barner, E. Gomez-Luna, Nonintrusive load monitoring based on switching voltage transients and wavelet transforms, in: *2012 Future of Instrumentation International Workshop (FIIW) Proceedings, 2012*, pp. 1–4. doi:10.1109/FIIW.2012.6378333.
- [69] Y. C. Su, K. L. Lian, H. H. Chang, Feature selection of non-intrusive load monitoring system using STFT and wavelet transform, in: *2011 IEEE 8th International Conference on e-Business Engineering, 2011*, pp. 293–298. doi:10.1109/ICEBE.2011.49.
- [70] H.-H. Chang, Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses, *Energie* 5 (11) (2012) 4569–4589. doi:10.3390/en5114569.
- [71] M. Gray, W. G. Morsi, Application of wavelet-based classification in nonintrusive load monitoring, in: *2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE), 2015*, pp. 41–45. doi: 10.1109/CCECE.2015.7129157.
- [72] S. M. Tabatabaei, S. Dick, W. Xu, Toward non-intrusive load monitoring via multi-label classification, *IEEE Transactions on Smart Grid* 8 (1) (2017) 26–40. doi:10.1109/TSG.2016.2584581.
- [73] J. M. Gillis, S. M. Alshareef, W. G. Morsi, Nonintrusive load monitoring using wavelet design and machine learning, *IEEE Transactions on Smart Grid* 7 (1) (2016) 320–328. doi:10.1109/TSG.2015.2428706.
- [74] H. Y. Lam, G. S. K. Fung, W. K. Lee, A novel method to construct taxonomy electrical appliances based on load signatures, *IEEE Transactions on Consumer Electronics* 53 (2) (2007) 653–660. doi:10.1109/TCE.2007. 381742.

- [75] T. Hassan, F. Javed, N. Arshad, An empirical investigation of v-i trajectory based load signatures for non-intrusive load monitoring, *IEEE Transactions on Smart Grid* 5 (2) (2014) 870–878. doi:10.1109/TSG.2013.2271282.
- [76] L. Du, D. He, R. G. Harley, T. G. Habetler, Electric load classification by binary voltage current trajectory mapping, *IEEE Transactions on Smart Grid* 7 (1) (2016) 358–365. doi:10.1109/TSG.2015.2442225.
- [77] J. Gao, E. C. Kara, S. Giri, M. Bergs, A feasibility study of automated plug-load identification from high-frequency measurements, in: *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2015, pp. 220–224. doi:10.1109/GlobalSIP.2015.7418189.
- [78] S. Fryze, Wirk- blind und scheinleistung in elektrischen stromkreisen mit nichtsinusoidalem verlauf von strom und spannung, *Elektrotech. Z.* 53 (25) (1932) 596–599.
- [79] T. D. Huang, W. S. Wang, K. L. Lian, A new power signature for nonintrusive appliance load monitoring, *IEEE Transactions on Smart Grid* 6 (4) (2015) 1994–1995. doi:10.1109/TSG.2015.2415456.
- [80] D. F. Teshome, T. D. Huang, K. L. Lian, Distinctive load feature extraction based on fryzes time-domain power theory, *IEEE Power and Energy Technology Systems Journal* 3 (2) (2016) 60–70. doi:10.1109/JPETS.2016.2559507.
- [81] S. Gupta, M. Reynolds, and S. Patel, “Electrisense: single-point sensing using EMI for electrical event detection and classification in the home,” in *Proceedings of the international conference on Ubiquitous computing*. ACM, 2010, pp. 139–148.
- [82] H. Kim, M. Marwah, M. Arlitt, G. Lyon, J. Han, Unsupervised disaggregation of low frequency power measurements, in: *Proceedings of the 2011 SIAM International Conference on Data Mining*, 2011, pp. 747–758. doi:10.1137/1.9781611972818.64.

- [83] Z. Wang, G. Zheng, Residential appliances identification and monitoring by a nonintrusive method, *IEEE Transactions on Smart Grid* 3 (1) (2012) 80–92. doi:10.1109/TSG.2011.2163950.
- [84] G. C. Koutitas, L. Tassioulas, Low cost disaggregation of smart meter sensor data, *IEEE Sensors Journal* 16 (6) (2016) 1665–1673. doi:10.1109/JSEN.2015.2501422.
- [85] Solène Kojtych. Introduction au traitement du signal et à l'analyse fréquentielle par transformées de Fourier. [Rapport Technique] École polytechnique de Montréal. 2019. hal-02193929ff.
- [86] M. Misiti, Y. Misiti, G. Oppenheim, and J-M Poggi. Wavelets Toolbox Users Guide. The MathWorks, 2000. Wavelet Toolbox, for use with MATLAB.
- [87] P.S. Addison. The Illustrated Wavelet Transform Handbook. IOP Publishing Ltd, 2002. ISBN 0-7503-0692-0.
- [88] Morlet, J. Wave propagation and sampling theory:Part II, Sampling theory and complex waves. 1982, *Geophysics*, 47(1), pp. pp 222-236.
- [89] MEZIANI Fadia, " Analyse du degré de sévérité pathologique des signaux phonocardiogrammes (PCGs) par application des transformées d'ondelettes" thèse de doctorat , université Abou Bekr Belkaid Tlemcen, 2013.
- [90] M.G.E. Schneiders. Wavelets in control engineering. Master's thesis, Eindhoven University of Technology, August 2001. DCT nr. 2001.38.
- [91] I. Daubechies. Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics, 1992. ISBN 0-89871-274-2.
- [92] S. MALLAT - A wavelet tour of signal processing - Book, 2nd Edition, Academic Press, 1999.
- [93] Hacine-Gharbi, A., Ravier, P., Harba, R. and Mohamadi, T. (2012) 'Low bias histogram-based, estimation of mutual information for feature selection', *Pattern Recognition Letters*, Vol. 33, No. 10, pp.1302–1308.

- [94] Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. *J Mach Learn Res* 3:1157–1182.
- [95] Kohavi R, John GH (1997) Wrappers for feature subset selection. *Artif Intell* 97(12):273–324.
- [96] L Yu and H. Liu, "Efficient Feature Selection via Analysis of Relevance and Redundancy," *J. Mach. Learn. Res.*, vol. 5, pp. 1205–1224, 2004.
- [97] S. Alelyani, J. Tang, and H. Liu, "Feature Selection for Clustering: A Review," in: C. Aggarwal and C. Reddy (eds.), *Data Clustering: Algorithms and Applications*, CRC Press, 2013.
- [98] Fleuret F, Guyon I (2004) Fast binary feature selection with conditional mutual information. *J Mach Learn Res* 5:1531–1555.
- [99] Guyon I, Elisseeff A (2006) An introduction to feature extraction. In: *Feature extraction, foundations and applications, studies in fuzziness and soft computing*, vol 207, Springer, Berlin, pp 1–25.
- [100] Kira, K. and Rendell, L.A., The feature selection problem: Traditional methods and a new algorithm. In: *Proceedings of Ninth National Conference on Artificial Intelligence*, 129-134, 1992.
- [101] Narendra, P.M. and Fukunaga, K., A branch and bound algorithm for feature selection. *IEEE Transactions on Computers*, C-26(9):917-922, September 1977.
- [102] Koller D, Sahami M. Toward optimal feature selection. In: *Proceedings of International Conference on Machine Learning*, 1996.
- [103] Kohavi, R. and Sommetfield, D., Feature subset selection using the wrapper method: Overfitting and dynamic search space topology. In: *Proceedings of First International Conference on Knowledge Discovery and Data Mining*, Morgan Kaufmann, 192-197, 1995.
- [104] M. Dash , H. Liu , Feature Selection for Classification, *Intelligent Data Analysis* 1(1997) 131-156

- [105] Duch W, Winiarski T, Biesiada J, Kachel A (2003) Feature selection and ranking filter. In: International conference on artificial neural networks (ICANN) and International conference on neural information processing (ICONIP), pp 251–254.
- [106] Weston J, Mukherjee S, Chapelle O, Pontil M, Poggio T, Vapnik V (2000) Feature selection for svms. In: Advances in neural information processing systems 13, MIT Press, pp 668–674.
- [107] H. Liu and H. Motoda, Feature Selection for Knowledge Discovery and Data Mining, London: Kluwer Academic Publishers, 1998.
- [108] Z. Xu, I. King, and M. R.-T. Lyu, "Discriminative Semi Supervised Feature Selection Via Manifold Regularization," IEEE Trans. Neural Networks, vol. 21, no. 7, pp. 1033– 1047, 2010.
- [109] A. Jović, K. Brkić and N. Bogunović, "A review of feature selection methods with applications," 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, 2015, pp. 1200-1205, doi: 10.1109/MIPRO.2015.7160458.
- [110] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "MIFS-ND: A information based feature selection method", Expert Systems with Applications, vol. 41, issue 14, pp. 6371–6385, 2014.
- [111] I. H. Witten and E. Frank, Data mining: Practical machine learning tools and techniques, San Francisco CA, USA: Morgan Kaufmann, 2011.
- [112] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation Based Filter Solution," in: Proc. 20th International Conference on Machine Learning (ICML-2003), Washington DC, USA, AAAI Press, pp. 856–863, 2003.
- [113] H. Liu and R. Setiono, "A Probabilistic Approach to Feature Selection-A Filter Solution," in: Proc. 13th International Conference on Machine Learning (ICML-1996), Bary, Italy, Morgan Kaufmann, pp. 319–327, 1996.

- [114] J. Tang, S. Alelyani, and H. Liu, "Feature Selection for Classification: A Review," in: C. Aggarwal (ed.), *Data Classification: Algorithms and Applications*. CRC Press, 2014.
- [115] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, Wiley-interscience, 2012.
- [116] M. R. Sikonja and I. Kononenko, "Theoretical and empirical analysis of Relief and Relief F," *Mach. Learn.*, vol. 53, pp. 23–69, 2003.
- [117] D. M. Witten and R. Tibshirani, "A framework for feature selection in clustering," *Journal of the American Statistical Association*, vol. 105, no. 490, pp. 713–726, 2010.
- [118] Y. Li, M. Dong, and J. Hua. "Localized feature selection for clustering," *Pattern Recognition Letters*, vol. 29, no. 1, pp. 10–18, 2008.
- [119] D. S. Modha and W.S. Spangler, "Feature weighting in k-means clustering," *Mach. Learn.*, vol. 52, no. 3, pp. 217–237, 2003.
- [120] M. Dash and Y.-S. Ong, "RELIEF-C: Efficient Feature Selection for Clustering over Noisy Data," in: *Proc. 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, Roca Raton, Florida, USA, pp. 869–872, 2011
- [121] P. S. Bradley and O. L. Mangasarian, "Feature selection via concave minimization and support vector machines," in: *Proc. 15th International Conference on Machine Learning (ICML- 1998)*, Madison, Wisconsin, USA, Morgan Kaufmann, pp. 82–90, 1998.
- [122] S. Maldonado, R. Weber, and F. Famili, "Feature selection for high-dimensional class imbalanced data sets using Support Vector Machines," *Information Sciences*, vol. 286, pp. 228–246, 2014.
- [123] Y. S. Kim, W. N. Street, and F. Menczer, "Evolutionary model selection in unsupervised learning," *Intelligent Data Analysis*, vol. 6, no. 6, pp. 531–556, 2002.

- [124] Hacine-Gharbi, A., Petit, M., Ravier, P. and Nemo, F. (2015) 'Prosody based automatic classification of the uses of French 'oui' as convinced or unconvinced uses', in International Conference on Pattern Recognition Applications and Methods (ICPRAM), ISBN: 978-989-758-077-2, pp.349–354.
- [125] M. Sandri and P. Zuccolotto, "Variable Selection Using Random Forests," in: S. Zani, A. Cerioli, M. Riani, and M. Vichi (eds.), Data Analysis, Classification and the Forward Search, Studies in Classification, Data Analysis, and Knowledge Organization, Springer, pp. 263–270, 2006.
- [126] G. C. Cawley, N. L. C. Talbot, and M. Girolami, "Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation," in: B. Schölkopf, J. C. Platt, and T. Hoffmann (eds.), Advances in Neural Information Processing Systems, MIT Press, pp. 209–216, 2007.
- [127] S. Das, "Filters, Wrappers and a boosting-based hybrid for feature selection," in: Proc. 18th International Conference on Machine Learning (ICML-2001), San Francisco, CA, USA, Morgan Kaufmann, pp. 74–81, 2001.
- [128] I. S. Oh, J. S. Lee, and B. R. Moon, "Hybrid genetic algorithms for feature selection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 11, pp. 1424–1437, 2004.
- [129] S. I. Ali and W. Shahzad, "A Feature Subset Selection Method based on Conditional Mutual Information and Ant Colony Optimization," International Journal of Computer Applications, vol. 60, no. 11, pp. 5–10, 2012.
- [130] S. Sarafrazi and H. Nezamabadi-pour, "Facing the classification of binary problems with a GSA-SVM hybrid system," Mathematical and Computer Modelling, vol. 57, issues 1-2, pp. 270–278, 2013.
- [131] Huang S, Dimensionality reduction in automatic knowledge acquisition: a simple greedy search approach. IEEE Trans Knowl Data Eng 2003, 15(6):1364–1373.
- [132] Mo D, Huang SH (2011) Feature selection based on inference correlation. Intell Data Anal 15(3):375–398.

- [133] Estévez PA, Caballero R (1998) A niching genetic algorithm for selecting features for neural networks classifiers. In: Perspectives in neural computation, Springer, New York, pp 311–316.
- [134] Mo D, Huang SH, Fractal-based intrinsic dimension estimation and its application in dimensionality reduction. IEEE Trans Knowl Data Eng, 2012, 24(1):59–71.
- [135] Sebban M, Nock R, A hybrid filter/wrapper approach of feature selection using information theory. Pattern Recognition , 2002, 35(4): 835–846.
- [136] Cover TM, Thomas JA, Elements of Information Theory. 2nd edn. Wiley Interscience, New Jersey, 2006.
- [137] Kullback S , Information theory and statistics. 2nd edn. Dover, 1997, New York.
- [138] Kullback S, Leibler RA, On information and sufficiency. Ann Math Stat 22:49–86, 1951
- [139] Battiti R. Using mutual information for selecting features in supervised neural net learning. IEEE Trans Neural Netw 5(4):537–550, 1994.
- [140] Lin D, Tang X. Conditional infomax learning: An integrated framework for feature extraction and fusion. ECCV (1), Leonardis A, Bischof H, Pinz A (eds.), Lecture Notes in Computer Science, Springer; 68-82.
- [141] Yang HH, Moody J. Data visualization and feature selection: New algorithms for non Gaussian data. in Advances in Neural Information Processing Systems, MIT Press, 1999; 687-693.
- [142] Bennasar M, Hicks Y, Setchi R. Feature selection using joint mutual information maximization. Expert Syst Appl 2015; 42(22):8520-8532.
- [143] Principe JC. Information theoretic learning: Renyi's entropy and kernel perspectives. 1st edn. Springer Publishing Company, Berlin, 2010.
- [144] Kraskov A, Stgbauer H, Grassberger P. Estimating mutual information. Phys Rev E 69:066,138, 2004.

- [145] J. R. Vergara and P. A. Estévez, "A review of feature selection methods based on mutual information," *Neural Computer. Appl.*, vol. 24, no. 1, pp. 175–186, 2014.
- [146] Hacine-Gharbi, A., Ravier, P. and Nemo, F. "Local and Global Feature Selection for Prosodic Classification of the Word's Uses.", In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2017)*, pages 711-717 , DOI: 10.5220/0006251407110717.
- [147] T. Drügman, M. Gurban, J.-P. Thirian, Relevant Feature Selection for Audio-visual Speech recognition, *International Workshop on Multimedia Signal Processing (MMSP)*, Chania, Crete, Greece, October 1-3, 2007.
- [148] Hacine-Gharbi, A., Ravier, P. and Mohamadi, T. (2009) 'Une nouvelle méthode de sélection des paramètres pertinents: application en reconnaissance de la parole', *Conférence TAIMA, Hammamet, Tunisie*.
- [149] Tsujishita T. On triple mutual information. *Adv Appl Math* 1995; 16(3): 269- 274.
- [150] A. Hacine-Gharbi, P. Ravier, R. Harba and T. Mohamadi . Combinaison et sélection des paramètres acoustiques pertinents par l'algorithme TMI. *Conférence : Traitement et Analyse de l'Information : Méthodes et Applications TAIMA'2013*, At: Hammamet – Tunisie, 2013
- [151] Bell AJ. The co-information lattice. *Analysis* pp 921–926. 2003.
- [152] Jakulin A, Bratko I (2003) Quantifying and visualizing attribute interactions. *CoRR cs.AI/0308002*, <http://arxiv.org/abs/cs.AI/0308002>.
- [153] McGill W. Multivariate information transmission. *Psychometrika* 19(2):97–116, 1954.
- [154] Huang JJ, Lv N, Li SQ, Cai YZ. Feature selection for classificatory analysis based on information theoretic criteria. *Acta Automatica Sinica* 2008; 34(3):383{392.
- [155] Moddemeijer, R. On estimation entropy and mutual information of continuous distributions. *s.l. : Signal Processing*, 1989 vol 3, pp. 233-248.

- [156] Joe, H. Estimation of entropy and other functionals of a multivariate density. *J. : Ann. Inst. Statist.* no. 4, pp. 683–697, Math, 1989, Vol 41.
- [157] . M. Ait Kerroum, A. Hammouch, and D. Aboutajdine. Textural feature selection by joint mutual information based on Gaussian mixture model for multispectral image classification. *J. : Pattern Recognition Letters*, no. 10, pp. 1168-1174, s. July 2010.
- [158] Sturges, H. The choice of a class-interval. , s.l. : *J. Amer. Statist. Assoc*, Vol.21, pp. 65–66, 1926.
- [159] Scott, D.W. On optimal and data-based histograms,". s.l. : *Biometrika*, Vol. 66, no. 3, pp 605–610, 1979,
- [160] , D. Freedman and P. Diaconis . On the Maximum Deviation Between the Histogram and the Underlying Density,". s.l. : *Zeitschrift fur Wahrscheinlichkeits theorie und verwandte Gebiete*, Vol 58, no. 2, pp. 139-167, , 1981.
- [161] C. Pascoal, M. R. Oliveira, A. Pacheco, R. Valadas, Theoretical evaluation of feature selection methods based on mutual information, *Neurocomputing* doi:<http://dx.doi.org/10.1016/j.neucom.2016.11.047>.URL <http://www.sciencedirect.com/science/article/pii/S092523121631462X>
- [162] Duch W. Feature extraction: foundations and applications. In: Guyon I, Gunn S, Nikravesh M, Zadeh L (editors). *Studies in Fuzziness & Soft Computing*. Berlin, Germany: Springer, 2006, pp. 89–117..
- [163] Bing Xue , " Particle Swarm Optimization for Feature Selection in Classification", thesis for the degree of Doctor of Philosophy in Computer Science, Victoria University of Wellington 2014.
- [164] Peng H, Long F, Ding C (2005) Feature selection based on mutual information criteria of max dependency, maxrelevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 27(8):1226–1238.

- [165] Gao J, Giri S, Kara EC, Berges M. PLAID: A public dataset of high-resolution electrical appliance measurements for load identification research: demo abstract. In: Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings; New York, NY, USA; 2014. pp. 198–199.
- [166] Grofman B, Owen G, Feld SL. Thirteen theorems in search of truth. *Theory and Decision* 1983; 15: 261-278.
- [167] Xu L, Krzyzak A, Suen CY. Method of combining multiple classifiers and their application to handwriting recognition. *IEEE Transactions on Systems, Man, and Cybernetics* 1992; 22 (3): 418-435.
- [168] Ho TK, Hull JJ, Srihari SN. Decision combination in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1994; 16 (1): 66-75.
- [169] Lam L, Suen CY. Application of majority voting to pattern recognition: an analysis of the behavior and performance. *IEEE Transactions on Systems, Man, and Cybernetics* 1997; 27 (5): 553-567.
- [170] Zou J, Nagy G. A comparative study of local matching approach for face recognition. *IEEE Transactions on Image Processing* 2007; 16 (10): 2617-2628.
- [171] Ogawara K, Fukutomi M, Uchida S, Feng Y. A voting-based sequential pattern recognition method. *PLoS ONE* 2013; 8 (10): e76980. doi: 10.1371/journal.pone.0076980.
- [172] Batra N, Parson O, Berges M, Singh A, Rogers A. A comparison of nonintrusive load monitoring methods for commercial and residential buildings 2014. arXiv preprint, arXiv: 1408.6595

