وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR- ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR ANNABA

جامعة باجي مختار- عنابة

Année : 2017

Faculté: Sciences de l'Ingéniorat
Département: Electronique

MEMOIRE
Présenté en vue de l'obtention du diplôme de : MASTER

Détection d'intrusion et d'événement anormale pour les systèmes SCADA

Domaine : Sciences et Technologie
Filière : Electronique
Spécialité : Automatique industriel

Par : EL BIR Soheib

DEVANT Le JURY

Président	**: REDJATI A/Ghani**	**MCB**	**Université Badji Moktar Annaba**
Directeur de mémoire	**: RAMDANI M.**	**Pr**	**Université Badji Moktar Annaba**
Examinateur	**: FEZARI M.**	**Pr**	**Université Badji Moktar Annaba**

ملخص

كشف الجدة هو تحديد بيانات جديدة أو غير معروفة أو إشارة حيث أن نظام آلة التعلم ليس على علم عند التشكيل.الكشف عن الجدة هو أحد المتطلبات الأساسية لنظام جيد للتصنيف أو تحديد الهوية، لأن في بعض الأحيان تحتوي معطيات الاختبار على بيانات من المعلومات عن الأشياء التي لم تكن معروفة في ذلك الوقت لتشكيل النموذج. إحدى الطرق الأكثر شعبية هو تحليل المكون الرئيسي (PCA) كأداة تحليل البيانات التي تقسيم مساحة القياس في فضاء جزئي للمكون الرئيسي الذي تحدث حينها تغيرات طبيعية وفضاء جزئي متبقي اين يمكن أن تتشكل الأخطاء.لعملية الرصد، إحصاءيتان اثنتان وهما ، SPE وT2، ويستخدمان للكشف عن التشوهات. ومع ذلك، لتحديد المتغيرات الخاطئة، خوارزميات التشخيص مثل مؤامرات المساهمة تواجه دائما مشاكل عند تطبيقها في عملية معقدة حقيقية.يتم إنجاز إعادة إعمار الخطأ عن طريق تحريك شعاع العينة أقرب ما يمكن إلى المساحة الفرعية للمكون الرئيسي. عندما يفترض خطأ فعلي. يتم التوصل إلى التقليل من الخطأ بالتنبؤ التربيعي (SPE). ويعرف رقم هوية الأخطاء من حيث المعاد بنائه. غير أن النظام الغير الخطي، PCA النوي يتعامل مع مكونات رئيسية أو الميزات التي ترتبط مع المتغيرات بطريقة غير خطية، للقيام بذلك نحسب المنتجات العددية في الفضاء المميز باستخدام دالة نوية في فضاء الإدخال. بإعتبار كل الخوارزميات التي يمكن التعبير عنها من قبل المنتجات العددية، وهذا القول دون استخدام صريحة من المتغيرات، الطريقة النواية تسمح لنا ببناء نسخة غير خطية من هذا.

الكلمات الرئيسية : فضاء جزئي,إعادة البناء,النواة,تحليل المكونات الرئيسي (ACP),مؤشرات الكشف,حدود الرقابة.

.

ABSTRACT

Abstract: Novelty detection is the identification of new or unknown data or signal that a machine learning system is not aware of during training. Novelty detection is one of the fundamental requirements of a good classification or identification system since sometimes the test data contains information about objects that were not known at the time of training the model. One of its popular methods is the principal component analysis (PCA) as a data anlysis tool which partitions the measurement space into a principal component subspace where normal variation occurs, and a residual subspace that faults may occupy. For the monitoring process, two statistics, T2 and SPE, are used to detect abnormalities. However to identify the faulty variables, the existing diagnosis algorithms such as contribution plots still have some troubles when they applied in real complex processes. Fault reconstruction is accomplished by sliding the sample vector as close as possible to the principal component subspace. When the actual fault is assumed, the maximum reduction in the squared prediction error (SPE) is achieved. A fault identification index is defined in terms of the reconstructed SPE. As for non linear systems, The Kernel PCA is interested in principal components,or features, that are nonlinearly related to the input variables. To do this, we compute the dot products in the feature space by means of a kernel function in the input space. Given any algorithm that can be expressed by dot products, that is, without explicit use of the variables, the kernel method allows us to construct a nonlinear versions of it.

Keywords: *PCA, Kernel, subspace, detection indices, Control limits, Reconstruction. .*

RÉSUMÉ

Abstract: La détection de nouveauté est l'identification des données nouvelles ou inconnues ou d'un signal qu'un système d'apprentissage machine n'est pas au courant lors de la formation. La détection de la nouveauté est l'une des exigences fondamentales d'un bon système de classification ou d'identification, car parfois les données de test contiennent des informations sur les objets qui n'étaient pas connus au moment de la formation du modèle. L'une de ses méthodes les plus populaires est l'analyse des composants principaux (ACP) comme un outil d'analyse de données qui partitionne l'espace de mesure dans un sous-espace de composant principal où se produisent des variations normales et un sous-espace résiduel que les défauts peuvent occuper. Pour le processus de suivi, deux statistiques, T2 et SPE, sont utilisées pour détecter des anomalies. Cependant, pour identifier les variables défectueuses, les algorithmes de diagnostic existants tels que les parcelles de contribution ont toujours des problèmes lorsqu'ils s'appliquent dans des processus complexes réels. La reconstruction des défauts s'effectue en glissant le vecteur échantillon aussi près que possible du sous-espace composant principal. Lorsque le défaut réel est supposé, la réduction maximale de l'erreur de prédiction au carré (SPE) est atteinte. Un indice d'identification des défauts est défini en termes de SPE reconstruit. Toutefois, pour les systèmes non linéaires, L'ACP à noyau s'intéresse aux composants principaux, ou aux fonctionnalités, qui sont liés de façon non linéaire aux variables d'entrée. Pour ce faire, nous calculons les produits scalaires dans l'espace caractéristique à l'aide d'une fonction à noyau dans l'espace d'entrée. Compte tenu de tout algorithme qui peut être exprimé par des produits scalaires, c'est à dire sans utilisation explicite des variables, la méthode à noyau nous permet de construire une version non linéaire de celle-ci.

Mots-Clés : *ACP, noyau, sous-espace , indices de détection, limites de contrôle, Reconstruction.*

REMERCIEMENTS

J'adresse mes remerciements a mes parents pour leur soutien inconditionnel, financier, moral, psychologique et matériel dont ils ont fait preuve depuis que j'ai commencé ma carrière d'études

Mes remerciements sont particulièrement à tous mes enseignants pour leur efforts durant mes années d'études et leur aide en toutes circonstances.

Je remercie mon encadreur Dr. Ramdani pour ses conseils, ses orientations, et son guide dans mon travail.

Je remercie vivement le président du jury et ses membres pour leur l'intérêt sur ma recherche et pour leur acceptation et propositions.

Mes remerciements s'adressent également à remercier toutes les personnes qui m'ont aidé dans la réalisation de ce mémoire dont je n'est pas cité leurs noms.

DÉDICACE

Je dédie ce mémoire :

à mes chers parents ,

à tous mes proches de la famille et particulièrement les frères et soeurs,

à tous mes chers amis et mes collègues de l'université,

à tous ceux et toutes celles , qui m'ont accompagné et soutenu durant ces années d'étude,

TABLE DES MATIÈRES

LISTE DES FIGURES

NOTATION

K	*matrice à Noyau.*
$\hat{X}$	*matrice de données estimée par le modèle ACP.*
$X \in \Re^{n \times m}$	*matrice de données.*
n	*nombre d'échantillons mesurées.*
m	*nombre de variables à surveiller.*
Λ	*matrice diagonale des valeurs propres.*
$P \in \Re^{m \times m}$	*matrice de vecteurs propres de la matrice de corrélation.*
$v_{i=1...m}$	*vecteurs variables (colonnes de la matrice de données).*
λ	*valeurs propres de la matrice de corrélation.*
$\vartheta_{i=1...n}$	*vecteurs observations (lignes de la matrice de données).*
$\bar{v}_j$	*moyenne de la jème variable.*
F	*l'espace caractéristique.*
$\vartheta_{i=1...n}$	*vecteurs observations (lignes de la matrice de données).*
σ_j^2	*variance de la jème variable.*
$\Sigma \in \Re^{m \times m}$	*matrice de covariance.*
Λ	*matrice diagonale des valeurs propres.*
$\tilde{P}$	$m - \ell$ *derniers vecteurs propres.*
x_{ij}	*échantillons du vecteur variable.*
$\hat{P}$	ℓ *premiers vecteurs propres.*
$\tilde{t}_j$	$m - \ell$ *dernières composantes principales jème variable.*
$T \in \Re^{n \times m}$	*matrice de composantes principales.*
$\ell < m$	*dimension de l'espace réduit (nb de CPs retenues).*
$\hat{C}$	*matrice représente le modèle ACP.*
$\eta_j^2(k)$	*indice de validité des capteurs.*
$\hat{t}_j$	ℓ *premières composantes principales de jème variable.*
ξ_j	*jème colonne de la matrice unitaire (direction de défauts).*
δ^2	*seuil de confiance de l'indicateur SPE.*
$\tilde{X}$	*matrice de données résiduelles.*
$\mathrm{T}^2_{\mathrm{lim}}$	*seuil de confiance de l'indicateur d'Hotelling.*
z_j	*reconstruction de la jème variable.*

C	*matrice de covariance en ACP à noyau.*
V^n	*l'enssemble des vecteurs propres en F.*
α	*les coordonnéees de v par rapport aux vecteurs caractéristiques.*
$\Phi(x)$	*l'image de x dans F.*
Φ	*matrice avec des colonnes données par la moyenne des données extraites dans F.*

IDS	*Intrusion Detection System.*
TCP	*Transmission Control Protocol.*
DR	*Taux de détection.*
FP	*Faux positive.*
ER	*Espace Résiduel.*
CPs	*Composantes Principales.*
EP	*Espace Principale.*
PCV	*Pourcentage Cumulé de la Variance.*
SVI	*Sensor Validity Index.*
PRESS	*Validation croisée.*
VER	*Variance de l'Erreur de Reconstruction.*
SPE	*Critère de l'Erreur quadratique de prédiction.*
T2	*Critère d'Hotelling .*
EWMA	*Exponentially Weighted Moving Average.*
DWDS	*Drinking Water Distribution System.*
SCADA	*Supervisory Control And Data Acquisition.*
KD99	*un jeu de données*

INTRODUCTION GÉNÉRALE

L'évolution actuelle des réseaux informatiques et particulièrement les réseaux SCADA ne présente pas que des avantages pour leurs utilisateurs, mais aussi des menaces et des risques. Surtout avec le déploiement parallèle des outils dédiés pour le piratage et l'utilisation malveillante. De ce fait, la protection et l'assurance du bon fonctionnement des réseaux sont devenues une vraie nécessité.

Pour cela, il existe plusieurs outils qui ont pour but d'assurer les trois impératifs, à savoir la confidentialité, l'intégrité et la disponibilité des données et des ressources dans système SCADA. Parmi ces outils, il y a les systèmes de détection d'intrusions IDS. Ces derniers permettent de détecter les intrusions dans un système SCADA et de prévenir le responsable des réseaux via des alertes.

De plus , un système de détection d'intrusion dans un système de contrôle industriel peut être vue comme un problème de classification qui consiste à définir des règles de classifications permettant de classer des objets dans des classes à partir d'attributs qualitatifs ou quantitatifs caractérisant ces objets dans le but de créer un modèle de décision qui est capable détecter ces intrusions, activités malveillantes ou comportement anormal qui affectent notre système, ces règles de classification sont définies par la sélection d'attributs qui précède souvent la classification supervisée (ou apprentissage supervisé) qui est principalement établi pour diminuer le nombre d'attributs à l'aide des algorithmes.

Une des méthodes statistiques les plus connues et appliquées dans différents domaines qui doivent détecter des anomalies dans leurs opérations régulières telles que la détection d'intrusions de réseau, le piratage, l'échec du moteur à réaction, l'apprentissage en machine et et beaucoup d'autres est enfin la détection de nouveauté .

Dans le cadre de notre mémoire , nous nous intéressons à la détection de nouveauté avec l'une de ses approches statistiques les plus connues celle de L'analyse en Composantes principales pour l'appliquer sur notre système de contrôle avancé SCADA celui de la distribution de l'eau potable pour maintenir la concentration du chlore dans le but de s'assurer du bon usage domestique à travers le réseau.

Ce mémoire de fin d'études se compose de quatres chapitres :

- *dans le premier chapitre nous passons en revue les approches les plus utilisées pour sécuriser les systèmes SCADA et montrer leur intérêt en citant quelques applications réels .*

- *dans le deuxième chapitre nous montrons qu'effectivement la détection d'intrusion en tant qu'approche de sécurité pour nos systèmes peut être considéré comme un problème de classification par plusieurs méthodes qui a pour but de construire un modèle de décision pour cette tache .*

- *dans le troisième chapitre nous présentons la détection de nouveauté pour pouvoir exposer l'une de ses méthodes les plus utilisés celle de L'analyse en Composantes principales dans les systèmes linéaires et non linéaires dont on aura besoin pour notre application .*

- *Le dernier chapitre sera consacré à la conception, la modélisation et l'implémentation de notre approche appliquée sur notre système de distribution d'eau potable avec des tests et résultats.*

CHAPITRE 1

SÉCURITÉ DES SYSTÈMES SCADA

Ce chapitre fournit les différents outils nécessaires pour sécuriser les systèmes SCADA en raison de leur vulnérabilité contre les attaques dangereuses qui peuvent endommager les systèmes et montrer quelques approches les plus utilisées pour cette tache qui n'est pas facile en montrant quelques exemples et applications réels et des techniques efficaces .

1.1 Système de détection d'intrusion pour SCADA

Contrôle de supervision et d'acquisition de données (SCADA) est un important système industriel contrôlé par ordinateur qui surveille et contrôle en permanence des sections différentes, infrastructures industrielles telles que les raffineries de pétrole, le traitement de l'eau et les systèmes de distribution et les installations de production d'énergie électrique, pour citer que quelques-uns. Un système SCADA est responsable de la surveillance et de contrôle des processus industriels et de l'infrastructure en recueillant des mesures et des données de commande provenant des appareil déployés au niveau du site. Les données collectées sont ensuite envoyées à un site central pour le processus et d'analyse. Les informations et les statuts sur les processus supervisés et suivi peuvent être affichés sur une interface homme-machine (HMI) à la station d'accueil de façon logique et organisée. Si un événement anormal se produit,

Un système de détection d'intrusion (IDS) est un matériel ou un logiciel autonome ou une combinaison de ceux-ci utilisés pour détecter des menaces qui pèsent sur les systèmes SCADA des deux attaques internes et externes, en surveillant et en analysant les activités sur un ordinateur hôte (host computer) ou un réseau. Une menace peut être considérée comme une activité malveillante visant à détruire la sécurité d'un système SCADA. Sous la menace, la confidentialité, l'intégrité ou la disponibilité de l'ordinateur hôte ou d'un réseau sont compromis. En outre, IDS

peut prévenir les menaces potentielles pour le système SCADA en détectant les précurseurs d'une attaque, l'accès non autorisé et opérations anormales,... etc. Selon l'emplacement et la source des données recueillies, dans l'informatique traditionnelle, IDSs peuvent être classés en network-based et host-based IDSs [1], Et cette catégorisation pourrait être similaire même aux systèmes SCADA. Toutefois, en raison de la nature différente des systèmes SCADA en termes d'architecture, les fonctionnalités, les appareils utilisés,... etc. SCADA IDS sont classés en fonction de la source des données recueillies : IDS basé sur les réseaux SCADA (SCADA Network-based) et IDS basé sur les applications (SCADA application-based).

1.1.1 IDS basé-réseaux SCADA

Un IDS basé-réseaux SCADA [2] capture les paquets de données qui sont communiqués entre des appareils tels que les points à points dans RTU / PLC, entre RTU / automates et CTU. Les appareils de surveillance sont toujours situés dans tout le réseau. Les informations contenues dans ces paquets de données capturées sont évaluées afin de déterminer si elles sont une menace ou pas. Si le paquet est suspect, les membres de l'équipe de sécurité seront alarmés pour une enquête plus approfondie. L'avantage des IDSs basés-réseaux SCADA est leur coût de calcul plus faible parce que seulement les informations contenues dans l'en-tête du paquet sont nécessaire pour le processus d'enquête, et donc un paquet de réseau SCADA peut être Scruté à la volée. Par conséquent, de grandes quantités de réseau peuvent être inspectées de manière satisfaisante et dans un délai acceptable [3].

Cependant, quand il y a un trafic réseau, un IDS basé-réseaux SCADA peut rencontrer des problèmes dans le suivi de tous les paquets et pourrait manquer une attaque lancée. La principale faiblesse est que le sens de fonctionnement du système SCADA contrôlé ne peut pas être déduit des informations fournies au niveau du réseau telles que l'adresse IP, le port TCP,... etc. Par conséquent, si la charge utile du paquet de réseau SCADA contient un message de contrôle malveillant, qui est conçu au niveau de l'application, les IDS basés-réseaux SCADA ne peuvent pas détecter si elle ne contraignent pas les spécifications du protocole utilisé ou le modèle de communication entre les périphériques du réseau SCADA [4].

1.1.2 IDS basé-applications SCADA

Les applications SCADA généralement connectent des informations précieuses sur les processus surveillées et contrôlées, et ce sont stockées dans des serveurs d'historien pour l'entretien, des fins commerciales, historiques et perspicacités. Les données SCADA, qui sont les données de mesure et de commande générés par des capteurs et des actionneurs, représentent la majorité de ces informations et, en outre, ils forment les informations opérationnelles pour un système SCADA donné à travers laquelle la présentation interne des systèmes contrôlés peut être déduite. Contrairement aux IDSs basés-éseaux SCADA qui inspectent uniquement des informations de niveau réseau, un IDS basé-application SCADA peut inspecter les données de haut niveau telles que les données SCADA pour détecter la présence d'un comportement inhabituel. Par exemple, les attaques de contrôle de haut niveau qui sont les menaces les plus difficiles à détecter par un IDS basé-réseaux SCADA [5], Peut être détectée en surveillant l'évolution des données SCADA puisque la source d'information d'un IDS basé-application SCADA peut être recueillie à partir des appareils du site à distance tels que PLC et RTU .

1.2 Les approches les plus utilisés d'IDS

1.2.1 Signature

Un SCADA IDS basé-signature détecte des activités malveillantes dans le trafic réseau SCADA ou les événements d'application en faisant correspondre les signatures d'attaques connues qui sont stockées dans une base de données spécifique. Le taux de faux positif dans ce type d'IDS est très faible et peut s'approcher de zéro. En outre, le temps de détection peut être rapide car il repose uniquement sur un processus de correspondance dans la phase de détection. Malgré les avantages énormes d'un IDS basé-signature, il ne pourra pas détecter une attaque inconnue dont la signature n'est pas connue ou qui n'existe pas dans sa base de données. Par conséquent, la base de données doit constamment être mise à jour avec des modèles de nouvelles attaques.

1.2.2 Anomalie

Un SCADA IDS basé-anomalie SCADA repose sur l'hypothèse que le comportement des activités intrusives est mathématiquement ou statistiquement différent des comportements normaux. C'est-à-dire qu'ils sont basés sur des techniques mathématiques ou statistiques avancées utilisées pour détecter le comportement anormal. Par exemple, le réseau SCADA normal peut être obtenu sur une période d'opérations "normale", puis une technique de modélisation est appliquée pour construire les fonctions normales du réseau SCADA. Dans la phase de détection, le degré d'écart entre le flux de réseau actuel Si le degré d'écart dépasse le seuil prééminent, le flux de réseau actuel sera marqué comme une activité intrusive. L'avantage principal de la détection basée sur l'anomalie par rapport à la détection basée sur la signature est que ces nouvelles (inconnues) attaques peuvent être détectées, même si elle est supérieure à un taux de faux positif élevé.

1.2.2.1 Type d'Anomalies avec quelques exemples concrets

Un aspect important d'une technique de détection d'anomalie est la nature de l'anomalie souhaitée. Les anomalies peuvent être classés en trois catégories suivantes :

Anomalie ponctuelle

Si une instance de données individuelles peut être considérée comme anormale par rapport au reste des données, l'instance est appelée comme une anomalie ponctuelle. Ceci est le plus simple type d'anomalie et est au centre de la majorité de la recherche sur la détection d'anomalies.

Par exemple, sur la figure (1.1), les points O_1 et O_2 ainsi que des points dans la région O_3 se situent en dehors des limites des régions normales, et sont donc des anomalies ponctuelles car ils sont différents des points de données normales.

A titre d'exemple de la vie réelle, pensez à la détection de la fraude par carte de crédit. Laissez-les ensemble de données correspondent aux transactions par carte de crédit d'un individu. Par souci de simplicité, supposons que les données sont définies en utilisant une seule fonction : montant dépensé. Une opération pour laquelle le montant dépensé est très élevé par rapport à la normale des dépenses

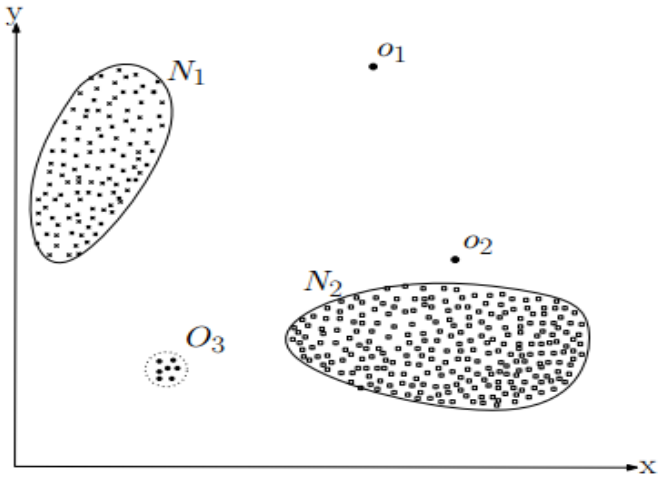


FIGURE *1.1 – Un simple exemple d'anomalies dans un jeu de données 2 dimensions.*

pour cette personne sera une anomalie ponctuelle.

Anomalies contextuelles

Si une instance de données est anormal dans un contexte spécifique (mais pas autrement), il est appelé comme une anomalie contextuelle (également appelée anomalie comme condition [18]. La notion de contexte est induite par la structure dans l'ensemble de données et doit être spécifiée comme une partie de la formulation du problème. Chaque instance de données est définie à l'aide après deux ensembles d'attributs :

- *attributs contextuels*

Les attributs contextuels sont utilisés pour déterminer le contexte (ou quartier) pour cette instance. Par exemple, dans des ensembles de données spatiales, la longitude et la latitude d'un emplacement sont les attributs contextuels. Dans les données séries chronologiques, le temps est un attribut contextuel qui détermine la position d'une instance sur la séquence entière.

- *attributs comportementaux*

Les attributs de comportement définissent les caractéristiques non-contextuelles d'une instance. Par exemple, dans un jeu de données spatiales décrivant la moyenne des précipitations du monde entier, la quantité de pluie à un endroit est un attribut de comportement.

Le comportement anormal est déterminée en utilisant les valeurs des attributs de comportement dans un contexte spécifique. Une instance de données pourrait être une anomalie contextuelle dans un contexte donné, mais une instance de données identiques (en termes d'attributs comportementaux) peuvent être considérés comme normaux dans un contexte différent. Cette propriété est essentielle pour identifier les attributs contextuels et comportementaux pour une technique de détection d'anomalie contextuelle.

Les anomalies contextuelles ont été les plus souvent explorées dans les données de séries chronologiques [19] et les données spatiales [20]. La figure (1.2) montre un exemple pour une série temporelle de température qui indique la température mensuelle d'une superficie de plus depuis quelques années. Une température de 35F peut être normale pendant l'hiver (à l'instant t_1) à cet endroit, mais la même valeur au cours de l'été (à temps t_2) serait une anomalie.

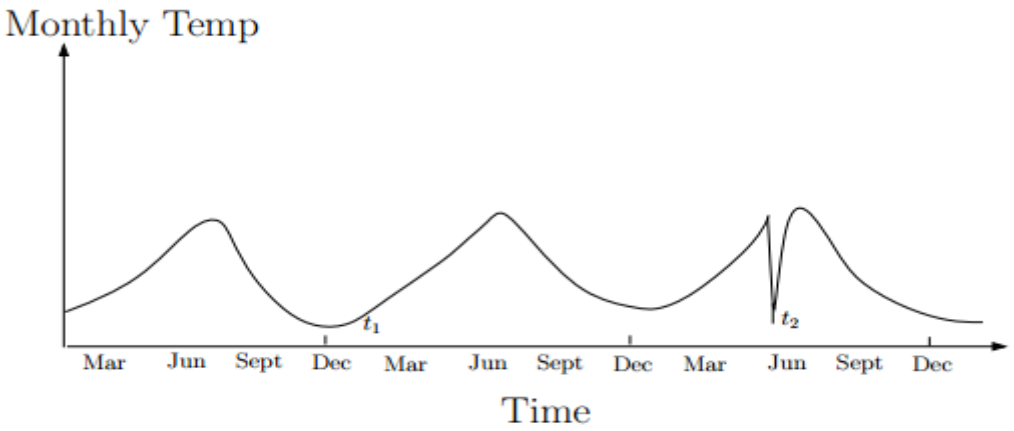


FIGURE *1.2 – anomalie contextuelle dans une série chronologique de la température.*

t_2 anomalie de contexte dans une série chronologique de la température. Notez que la température à l'instant t_1 est identique à celle au temps t_2, mais se produit dans un contexte différent et est donc pas considéré comme une anomalie. .

Anomalies collectives

Si une collection d'instances de données connexes est anormal par rapport à l'ensemble des données, il est appelé comme une anomalie collective. Les instances de données individuelles dans une anomalie collective ne peuvent pas être des anomalies par eux-mêmes, mais leur occurrence ensemble comme une collection

est anormale. La figure (1.3) illustre un exemple qui montre une sortie d'électrocardiogramme humaine [21]. La région en surbrillance indique une anomalie, car la même faible valeur existe pour un temps anormalement long (correspondant à une Contraction prématurée atriale). Notez que cette faible valeur par lui-même ne constitue pas une anomalie.

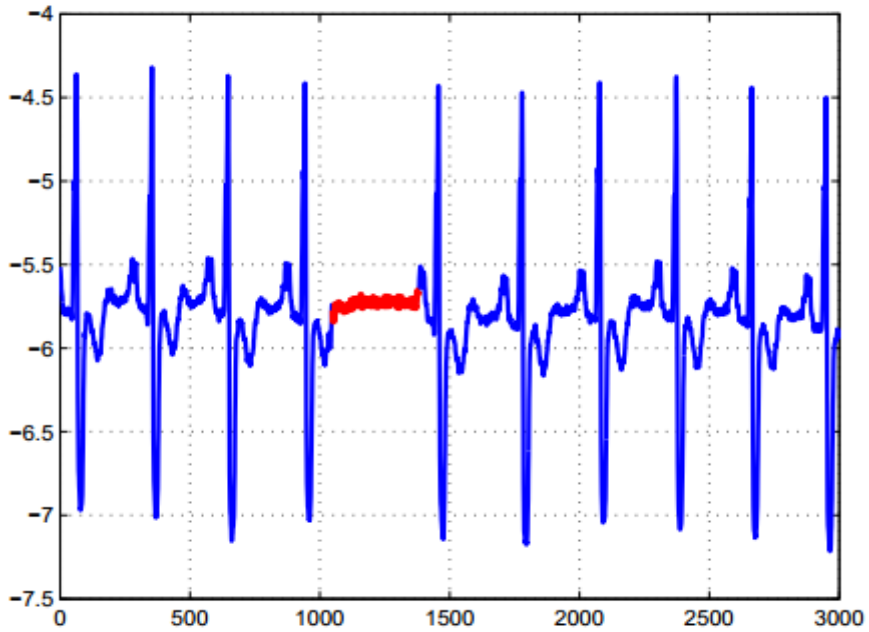


FIGURE *1.3 – anomalie collective correspondant à une Contraction prématurée atrial dans une sortie de l'électrocardiogramme.*

1.2.2.2 Les différentes techniques de détection d'anomalie avec des applications réelles

la détection d'anomalies a été le sujet d'un certain nombre d'enquêtes et d'articles de revue, ainsi que des livres. [8] fournissent une vaste enquête sur les techniques de détection des anomalies développées dans l'apprentissage de la machine et les domaines statistiques. Un examen général des techniques de détection des anomalies pour numériques ainsi que des données symboliques sont présentées par [9]. Un examen approfondi des techniques de détection de nouveauté à l'aide de réseaux de neurones et les approches statistiques a été présenté dans [10] et [11], respectivement. [12] présentent une étude des techniques de détection des anomalies

utilisées spécifiquement pour la détection cyber-intrusion. Une quantité importante de la recherche sur la détection des valeurs aberrantes a été faite dans les statistiques et a été examinée dans plusieurs livres [15] ; [16] ; [17], ainsi que d'autres articles de l'enquête [13] ; [14].

Le tableau 1.1 montre l'ensemble des techniques et des domaines d'application couverts et les divers articles d'enquête connexes mentionnés ci-dessus.

TABLE *1.1 – Techniques et applications avec des enquêtes :enquête 1*

		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
*6*Téchniques*	*Classification based*	●	●	●	●		●		
	Clustering based	●	●	●			●		
	Plus proche voisin	●	●	●			●		●
	Statistique	●	●	●		●	●	●	●
	Information théorique	●							
	Spectrale	●							
*7*Applications*	*Détection de cyber-intrusion*	●					●		
	Détection de fraude	●							
	Détection d'anomalies médicales	●							
	Détection de dommages industriels	●							
	Traitement d'image	●							
	Détection anomalie textuelle	●							
	Réseaux de capteurs	●							

1.2.3 Paramètre de concordance des motifs

La troisième approche de détection d'intrusion pour sécuriser les systèmes SCADA est plus subtile que les deux mentionnés précédemment. Cela explique le fait que les administrateurs système surveillent différents systèmes et attributs de réseau (ne ciblant pas nécessairement les problèmes de sécurité). En règle générale, les informations obtenues de cette manière ont un environnement spécifique constant. Cette méthode implique l'utilisation de l'expérience opérationnelle quotidienne des administrateurs comme base pour détecter les anomalies. Il peut être considéré comme un cas spécial de méthodes de profil normal. La différence réside dans le fait qu'un profil ici fait partie de la connaissance humaine.

C'est une technique très puissante, car elle permet des intrusions basées sur des attaques de type inconnu. L'opérateur du système peut détecter des modifica-

tions subtiles qui ne sont pas évidentes pour l'opérateur lui-même. Son inconvénient inhérent est lié au fait que les humains peuvent traiter et ne comprennent donc qu'une partie limitée de l'information à la fois, ce qui signifie que certaines attaques peuvent passer sans être détectées.

1.3 Conclusion

Nous avons enfin montré l'utilité et l'intérêt des approches les plus utilisées pour sécuriser les systèmes SCADA contre les activités malveillantes en désignant particulièrement la détection de l'anomalie avec ses méthodes, techniques et applications réelles vu qu'elle est parmi les thèmes les plus connus et traité par les chercheurs .

CHAPITRE 2

LA DÉTECTION D'INTRUSION DANS LES SYSTÈMES DE CONTRÔLE INDUSTRIEL PAR CLASSIFICATION

Dans ce chapitre, on va montrer que la détection d'intrusions dans les systemes de contrôle industriel peut être vue comme un problème de classification, en utilisant plusieurs méthodes sur un ensemble de données afin de construire un modèle et une règle de décision à l'aide de différentes techniques d'apprentissage dans le but de détecter les intrusions qui affectent nos systèmes de contrôle industriel.

2.1 La détection d'intrusions vue comme un problème de classification

Dans le travail qui va permettre de concevoir un système de détection d'intrusions, nous devons manipuler de grands volumes de données, et nous avons donc besoin d'en extraire de manière automatique ou semi-automatique des informations nécessaires et pertinentes en vue de leur utilisation opérationnelle. De ce fait, la détection d'intrusions peut être vue comme un problème de classification car elle s'intéresse à classifier des enregistrements en enregistrements sains et en tentatives d'intrusions.

Le cycle de vie d'une implémentation d'apprentissage automatique est le suivant [25] :

- *Obtention et nettoyage des données.*
- *Réalisation du modèle.*
- *Phase d'apprentissage.*
- *Phase de validation.*
- *Phase d'exécution.*

En ce sens on distingue deux type d'apprentissage automatique : apprentissage non supervisé associé à la classification non-supervisée et apprentissage supervisé

associé à la classification supervisée que nous utiliserons dans le cadre de ce projet de fin d'étude.

2.1.1 Classification non-supervisée :

La classification non supervisée ou le Clustering est une méthode statistique d'analyse de données, qui utilise les mesures de distances. Elle a pour objectif de grouper les données d'un ensemble en plusieurs sous-ensembles homogènes. Ces sous-ensembles contiennent des éléments qui ont des caractéristiques communes, c'est les critères de similarité. Le but des algorithmes de Clustering est donc de maximiser la distance interclasse et de minimiser la distance intra-classe pour avoir les sous-ensembles les plus distincts possibles [26].

2.1.2 Classification supervisée :

Dans l'étude des sciences cognitives [27], l'apprentissage supervisé est une technique d'apprentissage automatique (en anglais Machine Learning), qui consiste à faire apprendre à une machine une tâche déterminée. Pour cela, on doit constituer deux ensembles à partir de l'ensemble de données de départ, le premier pour faire l'apprentissage (généralement $\geq 60\%$ de l'ensemble), le second pour effectuer les tests et valider le procédé. Ce type d'algorithmes essaye de prédire avec le moins d'erreurs possibles le sous-ensemble auquel appartient chaque donnée testée. L'apprentissage supervisé s'intéresse principalement aux méthodes de régression de données (permet de prédire la valeur de sortie comme par exemple dans un ensemble potentiellement infini), et aux méthodes de classification de données (permet de prédire l'appartenance à une classe parmi un nombre fini de classes). Il existe de très nombreuses techniques d'apprentissage supervisé, nous citerons entre autres : Le modèle Naïve Bayes, Les machines à vecteur de support (en anglais Support Vector Machine SVM) et enfin Les arbres de décisions.

2.1.2.1 bayésien naïf

C'est une méthode qui utilise les règles de Bayes. Pendant la phase d'apprentissage, on calcule les probabilités des classes ainsi que celles de l'apparition

des attributs dans une classe .Pour dire qu'une instance appartient à une classe donnée, on calcule la probabilité d'apparition de chaque attribut dans cette classe. Un des principaux avantages de cette méthode est le besoin de faibles quantités d'informations pour la phase d'apprentissage et par conséquent une faible durée d'apprentissage.

2.1.2.2 Machines à vecteur de support

Les machines à vecteur support ou SVM (Support Vector Machine), sont des techniques récentes de classification supervisée, elles ont été introduites par Vladimir Vapnik, Bernhard Boser et Isabelle Guyon en 1992.

Ce sont des généralisations des classificateurs linéaires. Le but est séparer un espace grâce à un hyperplan. Dans le cas où on ne peut pas séparer les données par un hyperplan, on élève le rang des données non séparables à un espace de dimension supérieure en utilisant un noyau.

Les noyaux sont des fonctions qui associent à tout couple d'observations (x_i, x_j), une mesure calculée à travers leurs corrélations ou leur distances. L'intérêt des noyaux est de ramener le problème des données non linéairement séparables à un problème de données linéairement séparables [28].

2.1.2.3 Les arbres de décision

L'arbre de décision est un arbre acyclique où chaque nœud est étiqueté par un attribut, et chaque arête est étiquetée par un prédicat qui s'applique au nœud parent. Il s'agit de construire un arbre récursivement en choisissant l'attribut qui coupe au mieux les exemples dans leur classe [25], créant des nœuds pour chaque valeur de l'attribut choisi. Il existe plusieurs algorithmes de création d'arbres de décisions, nous citons : ID3, C4.5, CART. Afin de classifier un spécimen, on démarre de la racine et on descend jusqu'aux feuilles en respectant les prédicats.

C4.5 est une amélioration de l'algorithme ID3, ce dernier conçoit un arbre de décision d'une manière récursive en utilisant l'entropie de Shannon pour avoir l'attribut maximisant le gain d'informations. Cet algorithme utilise exclusivement des attributs discrets, d'où l'introduction de l'algorithme C4.5. Ce dernier a été

conçu en 1993 par Quinlan. Il permet de travailler à la fois avec des données discrètes et des données continues. Il peut aussi travailler avec des valeurs d'attributs absentes. C4.5 permet de compacter l'arbre de décision en éliminant les branches inutiles. [26].

2.2 Apprentissage supervisé pour la classification :

2.2.1 Apprentissage de la détection d'intrusion :

Ce formalisme appliqué sur un jeu de données KDD99 est censé proposer un algorithme d'apprentissage pour la classification dans le cadre de la détection d'intrusion avec ses résultats comparés au bayésien naïf et l'algorithme ID3 mais dans notre section on s'intéresse à montrer la procédure générale de la classification par l'apprentissage afin de construire un modèle de décision pour la détection d'intrusion qui est montrée dans des résultats expérimentaux par le taux de détection ,

la classification par l'apprentissage est constitué par :

- *La sélection des attributs de l'ensemble de données*
- *Construction des classificateurs*

la figure ci-dessous montre l'architecture générale d'un formalisme pour la selection des attributs et la classification :

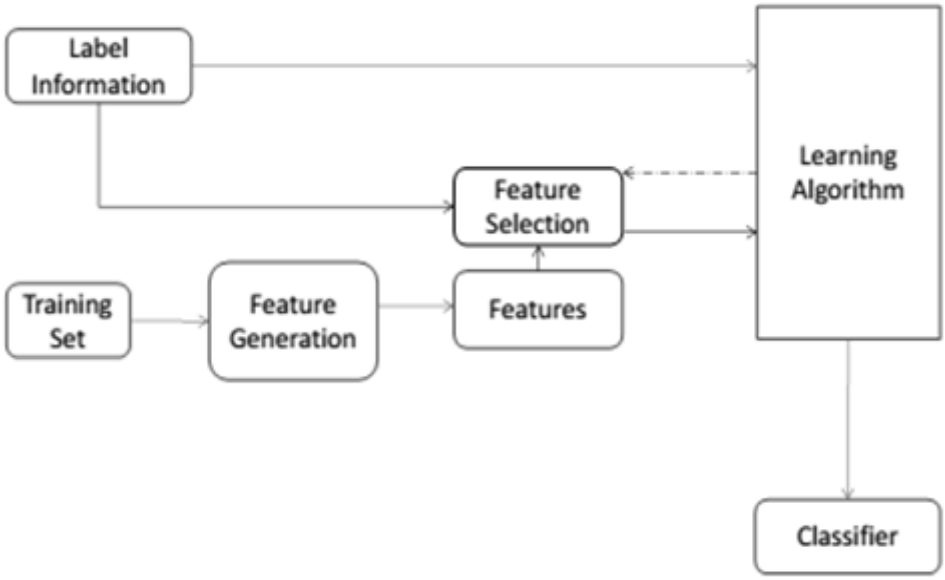


FIGURE *2.1 – Un formalisme générale de la sélection des attributs pour la classification*

La sélection des attributs de l'ensemble de données

La sélection efficace des attributs d'entrée à partir des jeux de données de détection d'intrusion est l'un des défis de recherche importants pour la construction d'IDS haute performance. Les attributs non pertinents et redondants de l'ensemble de données de détection d'intrusion peuvent conduire à un modèle complexe de détection d'intrusion ainsi qu'à réduire la précision de détection. Ce problème a été étudié lors des premiers travaux de [38], recherche sur l'ensemble de données de détection d'intrusion KDD99, où 41 attributs ont été construits pour chaque connexion réseau. Les méthodes de sélection d'attributs des algorithmes d'exploration de données identifient certains des attributs importants pour détecter des connexions réseau anormales. La sélection des attributs dans la détection d'intrusion à l'aide d'algorithmes d'exploration de données implique la sélection d'un sous-ensemble d'attributs d'un total d'attributs d'origine de l'ensemble de données, en fonction d'un principe d'optimisation donné. Trouver un sous-ensemble d'attributs utile est une forme de recherche. Idéalement, les méthodes de sélection des attributs recherchent les sous-ensembles d'attributs et essayent de trouver le meilleur

parmi les sous-ensembles candidats 2^N complets selon une fonction d'évaluation. Par conséquent, la construction d'IDS basée sur tous les attributs est impossible et la sélection des attributs devient très importante pour IDS.

Dans l'ensemble de données de détection d'intrusion KDD99, il existe 494021 exemples dans l'ensemble de données de formation de 10%. Le jeu de données KDD99 contient 22 types d'attaque différents qui peuvent être classés en quatre catégories principales : Déni de service (DoS), Remote to User (R2L), User to Root (U2R) et Probing. Il existe 41 attributs pour chaque connexion réseau qui possèdent des valeurs discrètes ou des valeurs continues. Les attributs dans l'ensemble de données KDD99 peuvent être divisés en trois groupes. Le premier groupe d'attributs est la fonctionnalité de base de la connexion réseau, qui comprend la durée, le prototype, le service, le nombre d'octets provenant des adresses IP source ou des adresses IP de destination et certains indicateurs dans les connexions TCP (des protocole de transport fiables). Le deuxième groupe d'attributs dans KDD99 se compose des fonctionnalités de contenu des connexions réseau et le troisième groupe est composé des caractéristiques statistiques qui sont calculées soit par une fenêtre temporelle, soit par une fenêtre de certains types de connexions. La sélection des attributs dans l'ensemble de données KDD99 a été largement utilisée comme méthode standard pour l'apprentissage par détection d'intrusion en réseau, et il a été constaté que les 41 attributs de l'ensemble de données KDD99 ne sont pas les meilleurs pour l'apprentissage par détection d'intrusion. Par conséquent, la performance d'IDS peut être encore améliorée en étudiant de nouvelles méthodes de sélection d'attributs [39].

Construction des classifieurs

La construction du classificateur est un autre défi de recherche important pour créer des IDS efficaces. De nos jours, de nombreux algorithmes d'exploration de données sont devenus très populaires pour classer les jeux de données de détection d'intrusion tels que l'arbre de décision, le classifieur bayésien naïf, le réseau neuronal, l'algorithme génétique et la machine vectorielle de support,.... etc. Cependant, la précision de la classification des algorithmes existants d'exploration de données doit être améliorée, car il est très difficile de détecter plusieurs nouvelles

attaques, vu que les attaquants modifient continuellement leurs modèles d'attaque. Les modèles de détection d'intrusion de réseau d'anomalie utilisent maintenant pour détecter de nouvelles attaques, mais les faux positifs sont généralement très élevés. La performance d'un modèle de détection d'intrusion dépend de ses taux de détection (DR) et de faux positifs (FP). DR est défini comme le nombre d'instances d'intrusion détectées par le système divisé par le nombre total d'instances d'intrusion présentes dans l'ensemble de données. FP est une alarme, qui s'élève pour quelque chose qui n'est pas vraiment une attaque. Il est préférable pour un modèle de détection d'intrusion de maximiser le DR et de minimiser le FP. Pour DR, nous pouvons modifier la fonction objective à 1-DR. Par conséquent, la construction du classifieur pour IDS est un autre défi technique dans le domaine de l'exploration de données.

2.2.2 Algorithme d'apprentissage proposé :

Compte tenu des données d'apprentissage $D = t_1, \ldots, t_n$ *où* $t_i = t_{i1}, \ldots, t_{ih}$ *et les données d'apprentissage D contient les attributs suivants :* $A_1, A_2, \ldots, A_n$ *et chaque attribut* A_i *contient les valeurs d'attribut suivantes* $A_{i1}, A_{i2}, \ldots, A_{ih}$. *Les valeurs d'attributs peuvent être discrètes ou continues. Aussi les données de formation D contient un ensemble de classes* $C = C_1, C_2, \ldots, C_m$. *Chaque exemple dans les données de formation D a une classe particulière* C_j. *L'algorithme calcule le gain d'information pour chacun des attributs* $A_1, A_2, \ldots, A_n$ *à partir des données de formation D.*

$$\mathrm{info(D)} = -\sum_{\mathrm{j=1}}^{\mathrm{m}} \frac{\mathrm{freq(c_i}, D)}{\mathrm{|D|}} \log_2 \left(\frac{\mathrm{freq(c_i}, D)}{\mathrm{|D|}} \right) \tag{2.1}$$

$$\inf(T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} \inf(T_i) \tag{2.2}$$

$$\text{Gain d'informations (Ai)} = \mathrm{Infos(D)} - \mathrm{Info(T)} \tag{2.3}$$

Ensuite, l'algorithme choisit l'un des meilleurs attributs A_i *parmi les attributs* $A_1, A_2, \ldots, A_n$ *à partir de données de formation D avec une valeur de gain d'infor-*

mation la plus élevée, Et divise les données de formation D en sous-ensembles de données D_1, D_2, ..., D_n en fonction des valeurs d'attribut choisies de A_i. L'algorithme estime alors les probabilités antérieures et conditionnelles pour chaque sous-ensemble de données $D_i = D_1$, D_2, ..., D_n et classe les exemples de sous-ensemble de données D_i en utilisant leurs probabilités respectives. La probabilité préalable P (C_j) pour chaque classe est évaluée en comptant combien de fois chaque classe se produit dans l'ensemble des données. Pour chaque attribut A_i du nombre d'occurrences de chaque valeur d'attribut A_{ij} peut être pris en compte pour déterminer P (A_i). De même, la probabilité conditionnelle P (A_{ij} | C_j) pour chaque valeur d'attribut A_{ij} peut être estimé en comptant la fréquence de chaque valeur dans la classe dans l'ensemble des données. Pour classer un exemple dans l'ensemble des données, les probabilités préalable et conditionnelles générées à partir de l'ensemble des données sont utilisées pour faire la prédiction. Cela se fait en combinant les effets des différentes valeurs d'attributs de l'exemple. Supposons que l'exemple e_i a des valeurs d'attributs indépendants A_{i1}, A_{i2}, ..., A_{ip}, nous savons P (A_{ik} | C_j), pour chaque classe C_j et attribut A_{ik}. Nous estimons alors P (e_i | C_j) par :

$$\mathrm{P(e_i|C_j) = P(C_j) \prod_{k=1 \to p} P(A_{ij}|C_j)} \tag{2.4}$$

Pour classer un exemple dans l'ensemble de données, l'algorithme estime la probabilité que e_i est dans chaque classe. La probabilité que e_i est dans une classe est le produit des probabilités conditionnelles pour chaque valeur d'attribut avec une probabilité antérieure pour cette classe. La probabilité postérieure P (C_j | e_i) est alors trouvée pour chaque classe et l'exemple classe avec la probabilité postérieure la plus élevée pour cet exemple. L'algorithme continuera jusqu'à ce que tous les exemples de subdatasets ou sous-subdatasets sont correctement classés. Lorsque l'algorithme classe correctement tous les exemples de tous les sous / sous-sous-ensembles de données, l'algorithme se termine et les probabilités antérieures et conditionnelles pour chaque sous / sous-sous-ensembles de données sont conservés pour la classification future des exemples invisibles. La procédure principale de l'algorithme proposé est décrit comme suit :

Algorithme

- *Entrée : Formation Dataset D*
- *Sortie : Modèle de détection d'intrusion*
- *Procédure :*
  - *1. Calculer le gain d'informations pour chaque attributs : $A_i = \{A_1, A_2, \cdots, A_n\}$ à partir des données de formation D en utilisant l'équation (2.3).*
  - *2. Choisir un attribut A_i à partir des données de formation D avec la valeur maximale de gain d'informations.*
  - *3. Diviser les données de formation D dans les sous-ensembles de données : $D_i = \{D_1, D_2, \cdots, D_n\}$ en fonction des valeurs d'attribut de A_i.*
  - *4. Calculer les probabilités P (C_j) antérieures et conditionnelles P (A_{ij} | C_j) de chaque sous-ensemble de données D_i.*
  - *5. Classifier les exemples de chaque sous-ensemble de données D_i avec leurs probabilités antérieures et conditionnelles respectives.*
  - *6. Si un exemple de sous-ensemble de données D_i est mal classé, calculer à nouveau le gain d'informations des attributs du sous-ensemble de données D_i, choisissez le meilleur attribut Ai avec la valeur de gain d'information maximale à partir du sous-ensemble de données D_i, Diviser le sous-ensemble de données D_i en sous-sous-ensembles de données D_{ij} et de nouveau calculer les probabilités antérieures et conditionnelles pour chaque sous-sous-ensemble de données D_{ij}. Enfin, classifiez les exemples de sous-sous-ensembles de données en utilisant leurs probabilités antérieures et conditionnelles respectives*
  - *7. Continuez ce processus jusqu'à ce que tous les exemples de sub / subsub-datasets soient correctement classés.*
  - *8. Préservé toutes les probabilités antérieures et conditionnelles pour chaque sous-ensemble de données D_i ou sous-sous-ensemble de données D_{ij} pour la classification future d'exemples non connus.*

TABLE *2.1 – Différents types d'attaques en KDD'99 dataset*

4 Main Attack Classes	*22 Attack Classes*
Probing	*ipsweep, nmap, portsweep, satan*
Denial of Service	*back, land, eptune, pod, smurf, teardrop*
User to Root	*buffer_overflow, perf, loadmodule, rootkit*
Remote to User	*ftp_write,guess_passwd,imap,multihop,phf,spy,warezclient/master*

2.2.3 Résultats expérimentaux :

Flux de données de la détection d'intrusion

L'expérience a été réalisée sur un flux de données réel appelé «ensemble de données de détection d'intrusion», qui a été utilisé dans la compétition Coupe 1999 de découverte et de minutage de données (KDD) [23] Dans l'ensemble de données KDD99, le flux de données d'entrée contient les détails des connexions réseau, tels que le type de protocole, la durée de la connexion, le type de connexion, etc. Chaque échantillon de données dans l'ensemble de données KDD99 représente la valeur d'attribut d'une classe dans le flux de données réseau et chaque classe est étiquetée Soit normalement, soit comme une attaque avec exactement un type d'attaque spécifique. Au total, 41 fonctionnalités ont été utilisées dans l'ensemble de données KDD99 et chaque connexion peut être classée en cinq classes (une classe normale et quatre classes d'intrusion principales : sonde, DOS, U2R, R2L). Il existe 22 types d'attaques différentes qui sont regroupées dans les quatres principaux types d'attaques (sondes, DOS, U2R, R2L) tabulées dans le Tableau 2.1. Une attaque par déni de service (en anglais Denial Of Services abrégé en DOS) est un type d'attaque visant à rendre indisponible, pendant un temps indéterminé, les services ou ressources d'une organisation, son principe est de saturer les serveurs et de les paralyser en leur envoyant des milliers de requêtes depuis des stations différentes, qui ne sont pas forcément situées dans le même emplacement.

Le paramètre expérimental est le même que celui utilisé dans la Coupe KDD99 [23] prenant 10% de l'ensemble du flux réel de données brutes (494021 échantillons de données) pour la formation et 311029 échantillons de données pour les tests. Le tableau 2.2 montre le nombre d'exemples de formation et de test pour chaque classe dans l'ensemble de données KDD99.

TABLE *2.2 – Nombre d'exemples en KDD'99 dataset*

Attack Types	*Training examples*	*Testing examples*
Normal	*97277*	*60592*
Probing	*4107*	*4166*
Denial of Service	*391458*	*237594*
User to Remote	*52*	*70*
Remore to User	*1126*	*8606*
Total examples	*494020*	*311028*

Analyse expérimentale :

Nos expériences ont deux étapes, à savoir apprendre et classer les données de formation, puis classer les données de test. Dans la première phase, les attributs importants des données de formation de KDD99 sont sélectionnés par des valeurs maximales de gain d'information, puis les probabilités antérieures et conditionnelles sont utilisées pour construire un modèle de détection à l'aide des attributs sélectionnés. Dans la deuxième phase, les données de test de KDD99 ont traversé le modèle qualifié pour détecter les intrus et trouver les taux de détection et les faux positifs du modèle de détection. Dans l'expérience, 41 attributs de l'ensemble de données KDD99 sont étiquetés dans l'ordre A_1, A_2, A_3 , A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}, A_{11}, A_{12}, A_{13}, A_{14}, A_{15}, A_{16}, A_{17}, A_{18}, A_{19}, A_{20}, A_{21}, A_{22}, A_{23}, A_{24}, A_{25}, A_{26}, A_{27}, A_{28}, A_{29}, A_{30}, A_{31}, A_{32}, A_{33}, A_{34}, A_{35}, A_{36}, A_{37}, A_{38}, A_{39}, A_{40}, A_{41}. *Toutes les expériences ont été effectuées à l'aide d'un processeur Intel Core 2 Duo processeur 2.0 GHz (2 Mo de cache, 800 MHz FSB) avec 1 Go de RAM. Nous avons sélectionné les attributs importants à partir du jeu de données KDD99 en utilisant notre algorithme proposé et nous avons découvert que 19 attributs sont importants et 22 attributs sont redondants ou moins importants. Les 19 attributs importants sont :* A_1, A_2, A_3 ,A_1, A_3, A_4, A_5, A_6, A_8, A_9, A_{10}, A_{11}, A_{13}, A_{15}, A_{16}, A_{17}, A_{18}, A_{19}, A_{23}, A_{24}, A_{32} *et* A_{33}. *De l'autre ct, les* 22 *attributs redondants sont* A_2, A_7, A_{12}, A_{14}, A_{20}, A_{21}, A_{22}, A_{25}, A_{26}, A_{27}, A_{28}, A_{29}, A_{30}, A_{31}, A_{34}, A_{35}, A_{36}, A_{37}, A_{38}, A_{39}, A_{40} , *Et* A_{41}. *Après avoir identifié les attributs importants à partir du jeu de données KDD99, nous avons calculé les probabilités antérieures et conditionnelles pour construire le modèle de détection d'intrusion. Ensuite, l'ensemble de données de test de KDD99 est utilisé sur le modèle de détection pour classer*

les exemples comme une attaque ou une normale. La comparaison de performance basée sur le taux de détection (DR) et les faux positifs (FP) entre 41 attributs et 19 attributs pour 5 classes d'attaque sur l'ensemble de données KDD99 en utilisant l'algorithme ID3, le classifieur bayésien naïf et notre algorithme proposé sont listés dans le Tableau 2.3, Tableau 2.4, Tableau 2.5 et Tableau 2.6.

TABLE *2.3 – Detection Rates (%) Using 41 Attributes*

Classes	*ID3 Algorithm*	*NB classifier*	*Proposed Algorithm*
Normal	*99.63*	*99.27*	*99.65*
Probe	*97.85*	*99.11*	*99.21*
DoS	*99.51*	*99.69*	*99.71*
U2R	*49.21*	*64.00*	*99.17*
R2L	*92.75*	*99.11*	*99.25*

TABLE *2.4 – Detection Rates (%) Using 19 Attributes*

Classes	*ID3 Algorithm*	*NB classifier*	*Proposed Algorithm*
Normal	*99.71*	*99.65*	*99.82*
Probe	*98.22*	*99.35*	*99.72*
DoS	*99.63*	*99.71*	*99.75*
U2R	*86.21*	*64.84*	*99.47*
R2L	*97.79*	*99.15*	*99.35*

TABLE *2.5 – False Psitives (%) Using 41 Attributes*

Classes	*ID3 Algorithm*	*NB classifier*	*Proposed Algorithm*
Normal	*0.10*	*0.08*	*0.07*
Probe	*0.55*	*0.45*	*0.42*
DoS	*0.04*	*0.04*	*0.04*
U2R	*0.14*	*0.14*	*0.12*
R2L	*10.10*	*8.02*	*7.87*

Par conséquent, il ressort clairement du résultat ci-dessus qu'une sélection significative d'attributs améliore les performances du modèle de détection

2.2.4 Conclusion

Dans ce formalisme, nous avons proposé une nouvelle approche d'apprentissage pour la détection d'intrusion réseau qui effectue une réduction de données

TABLE *2.6 – False Psitives (%) Using 19 Attributes*

Classes	*ID3 Algorithm*	*NB classifier*	*Proposed Algorithm*
Normal	*0.06*	*0.05*	*0.05*
Probe	*0.51*	*0.32*	*0.28*
DoS	*0.04*	*0.04*	*0.03*
U2R	*0.12*	*0.12*	*0.10*
R2L	*7.34*	*6.87*	*6.24*

en sélectionnant un sous-ensemble important d'attributs. La performance de notre approche proposée sur l'ensemble de données de détection d'intrusion KDD99 a atteint des taux de détection d'équilibre pour cinq classes. Il a également réduit les faux positifs par rapport à l'algorithme ID3 et au classifieur bayésien naïf. Les résultats expérimentaux montrent que la sélection significative d'attributs améliore les performances d'IDS. Les attaques de l'ensemble de données KDD99 détectées avec une précision de 99% en utilisant notre approche proposée. Le travail futur met l'accent sur l'amélioration des faux positifs de l'attaque R2L et applique le modèle de détection dans IDS du monde réel.

2.3 Conclusion

Dans ce chapitre nous avons enfin montrer qu'effectivement la détection d'intrusion peut etre considéré comme un problème de classification avec plusieurs méthodes et algorithmes d'apprentissage pour construire un modèle ou une régle de décision dans le but de détecter ces intrusions qui affectent nos systémes de contrôle industriel, elle consiste à établir la sélection des attributs dans un ensemble de données pour enfin passer à la construction des classifieurs.

CHAPITRE 3

DÉTECTION DE NOUVEAUTÉ

Dans ce chapitre, nous allons présenter La détection de la nouveauté ou anomalie en particulièr qui est présentée dans le chapitre 1 et avec le même principe , alors. La détection de la nouveauté qui détermine , à partir d'un ensemble de données, les points considérés comme normaux et ceux qui sont nouveaux. Compte tenu d'un ensemble de données, nous formons un ensemble d'observations normales et déterminons une distance de cette région normale qui classerait un point comme nouveau. Autrement dit, si une observation a une valeur supérieure à la valeur normale décidée, elle est considérée comme nouvelle. Il existe beaucoup d'approches pour la Détection de la nouveauté, y compris l'analyse en Composantes principales qui est parmi les méthodes statistiques les plus efficaces et connues et c'est ce que nous allons voir .

3.1 L'Analyse en Composantes principales ACP

L'Analyse en Composantes principales "ACP" est une des méthodes descriptives multidimensionnelles appelées méthodes factorielles utilisées pour la détection de nouveauté .Cette technique statistique permet d'étudier simultanément les relations qui existent entre les variables, et de réduire la dimension-alité d'un ensemble de données d'une taille importante afin d'analyser et traiter ces données.

L'ACP peut servir à mieux connaître les données sur lesquelles on travaille, à détecter éventuellement des valeurs suspectes, et si on parle de la classification des données, la réduction de l'espace de représentation permet de diminuer le temps de classification pour toute nouvelle information et éventuellement de déterminer le nombre de groupes à construire, elle peut être aussi une intermédiaire de calcule en vu d'analyse ultérieurs.

Cette section comprend deux parties, la première est consacrée à la modélisation par ACP linéaire et la deuxième consacrée à la détection et la localisation par le modèle obtenu.

Avant de d'approfondir, il faut tout d'abord bien comprendre quelques notions de statistique liées à cette approche :

3.1.1 Notions de base

Tableau de données :

L'ACP propose à partir d'un tableau (on dit aussi matrice) de données [voir figure (3.1)], une représentation géométrique permet de former ce que l'on appelle nuage de points, où chaque point est positionné dans un repère en fonction de ses coordonnées [voir figure (3.2)].

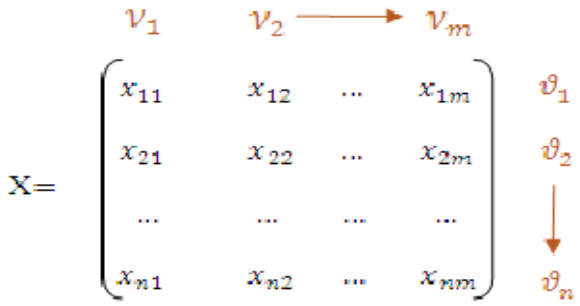


FIGURE *3.1 – Structure d'un tableau de données.*

Variable :

Chaque une des colonnes de la matrice de données est appelée variable, et chaque variable (ou colonne) est un vecteur contient "n" échantillons ou paramètres, qui sont les valeurs numérique de ce variable pour chaque une des observations.

Observation :

Chaque une des lignes de la matrice de données est appelée observation, c'est un vecteur (ligne) de "m" échantillons qui sont les valeurs numériques des m variables pour cette observation.

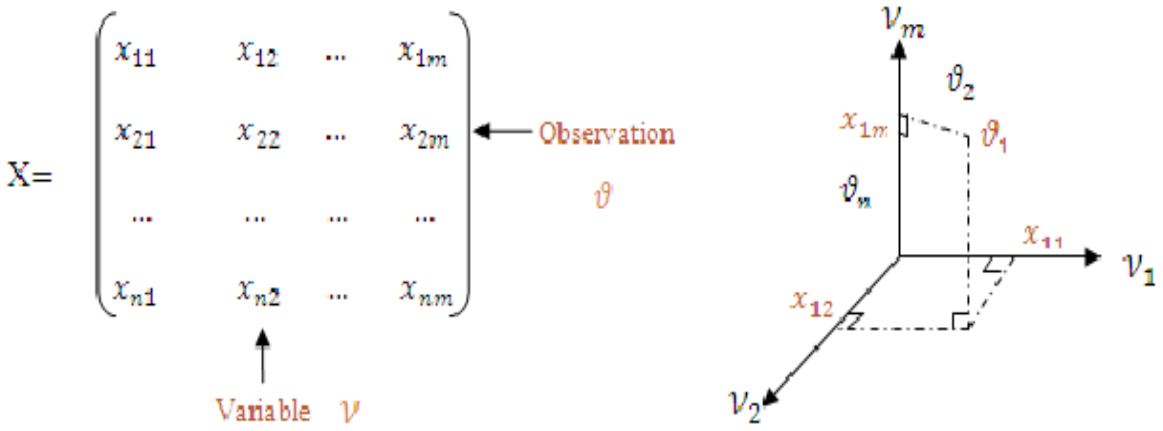


FIGURE *3.2 – Représentation géométrique des observations.*

Moyenne :

La moyenne d'une variable v où $v_j^T = \begin{bmatrix} x_{1j} & x_{2j} & \cdots & x_{nj} \end{bmatrix}$ *et* $j = \begin{bmatrix} 1 & \cdots & m \end{bmatrix}$

$$\bar{v}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{3.1}$$

Variance :

Cette grandeur est déterministe et toujours supérieurs ou égale à zéro, elle peut s'interpréter comme la mesure des fluctuations (la dispersion) de la variable aléatoire autour de sa moyenne.

si $j = \begin{bmatrix} 1 & \cdots & m \end{bmatrix}$ *la variance de chaque variable a la formule suivante :*

$$\sigma_j^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_{ij} - \bar{v}_j) \tag{3.2}$$

3.1.2 Modélisation par ACP Linéaire

Pour trouver un modèle basé sur l'ACP linéaire on a besoin d'une base de données contient des variables à surveiller pour un ensemble des mesures effectuées sur le fonctionnement normale du système : [voir figure (3.1)]

Normalisation de la matrice de données :

Premièrement on va chercher de centrer la matrice de données sur le centre de gravité du nuage de points, parce que le point 0 *correspondant au vecteur de coordonnées toutes nulles n'est pas forcément une origine satisfaisante, car si les coordonnées des points du nuage sont grandes, le nuage est éloigné de cette origine, alors il faut choisir une origine liée au nuage de points lui-même [voir figure (3.3)]. Le centre de gravité* G *du nuage des observations est alors le point dont les coordonnées sont les valeurs moyennes des variables :*

$$G = \left(\bar{v}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{i1} \quad \bar{v}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{i2} \quad \cdots \quad \bar{v}_m = \frac{1}{n}\sum_{i=1}^{n} x_{im} \right) \tag{3.3}$$

Où $\frac{1}{n}$ *est le poids des observations*

Et alors on va travailler sur le tableau des données centrées [voir figure (3.3)]

Deuxièmement la réduction de la matrice de données centrée par la transformation suivante :

$$X_j = \frac{X_C}{\sigma_j} \tag{3.4}$$

Où σ_j *est la variance de la jeme variable*

$$\mathrm{X_C} = \begin{pmatrix} x_{11} - \bar{v}_1 & \dots & x_{1j} - \bar{v}_2 & \dots & x_{1p} - \bar{v}_m \\ & & \vdots & & \\ x_{i1} - \bar{v}_1 & \dots & x_{ij} - \bar{v}_2 & \dots & x_{ip} - \bar{v}_m \\ & & \vdots & & \\ x_{n1} - \bar{v}_1 & \dots & x_{nj} - \bar{v}_2 & \dots & x_{np} - \bar{v}_m \end{pmatrix}$$

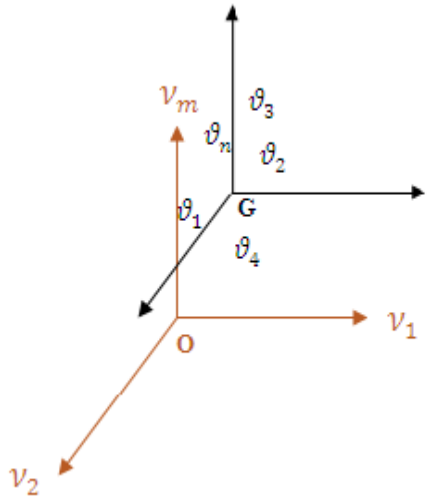


FIGURE *3.3 – Nuage centré sur le centre de gravité.*

La matrice de variance-covariance :

$$\Sigma = \left(\frac{1}{n} X^T X\right) \in \Re^{m \times m} \tag{3.5}$$

Axes principaux et composantes principales :

Notons $P = \begin{pmatrix} P_1 & \cdots & P_m \end{pmatrix} \in \Re^{m \times m}$ *matrice de vecteurs propres et les* λ_i *sont les valeurs propres de la matrice de variance-covariance où la décomposition en valeurs singulières de cette matrice donne :*

$$\Sigma = P \Lambda P^T \tag{3.6}$$

Où Λ *est la matrice diagonale des valeurs propres dans l'ordre décroissant :*

$$\Sigma = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_m \end{pmatrix}_{\lambda_1 > \lambda_2 > \cdots > \lambda_m} \tag{3.7}$$

Les composantes principales sont les vecteurs de projections des observations (vecteurs de mesures) sur les axes de direction des vecteurs propres de la matrice de corrélation des données, ces composantes sont obtenus par :

$$T = XP \tag{3.8}$$

$$T = \begin{bmatrix} T_1 & T_2 & \cdots & T_m \end{bmatrix} \in \Re^{n \times m}$$

Modèle ACP :

L'objectif est de diminuer la dimension de l'espace de représentation des observations, une fois déterminer le nombre $\ell < m$ *de composantes à retenir (la dimension réduite à choisir), la matrice* X *peut être approximer et pour cela, la*

matrice P est partitionnée sous la forme :

$$P = \begin{pmatrix} \hat{P}_\ell & \tilde{P}_{m-\ell} \end{pmatrix}$$

$\hat{P}_\ell$:les ℓ premiers vecteurs propres constituent l'espace principale EP $\tilde{P}_{m-\ell}$: les $(m-\ell)$ derniers vecteurs propres constituent l'espace résiduel ER

Et alors l'ACP détermine la transformation optimale suivante :

$$T = XP \Rightarrow \hat{T} = X\hat{P}$$

$$X = TP^T \Rightarrow \hat{X} = \hat{T}\hat{P}^T \Rightarrow \hat{X} = X\hat{P}\hat{P}^T \tag{3.9}$$

Notons que :

$$\hat{C} = \hat{P}\hat{P}^T \tag{3.10}$$

$$\left(I_m - \hat{C}\right) = \left(I_m - \hat{P}\hat{P}^T\right) = \tilde{C} \tag{3.11}$$

Et alors :

$$\hat{X} = X\hat{C} \tag{3.12}$$

et

$$\tilde{X} = X\tilde{C} \tag{3.13}$$

L'estimation de X par l'ACP est donnée par la formule suivante :

$$\hat{X} = \sum_{j=1}^{\ell} (T_j P_j^T)$$

Et alors les données restantes ou résiduelles de la matrice de données X représentent la perte de l'information acceptable de cette technique :

$$\tilde{X} = \sum_{j=\ell+1}^{m} (T_j P_j^T)$$

Où

$$X = \hat{X} + \tilde{X} \tag{3.14}$$

3.1.3 Choix de la dimension de l'espace réduit

Concernant le choix du nombre de composantes principales qui doit être retenu, il existe plusieurs critères tels que le pourcentage cumulé de la variance totale "PCV", la moyenne des valeurs propres, la validation croisée "PRESS" , variance de l'erreur de reconstruction "VER", et cette dernière approche avec la première représentent notre choix pour déterminer la dimension de l'espace principal dans cette étude :

3.1.3.1 Variance de l'erreur de reconstruction (VER) :

Cette technique basée sur la reconstruction des variables à partir du modèle et des autres variables de la matrice de données. Le principe est de chercher le nombre de composantes principales qui permet de minimiser la variance de l'erreur de reconstruction ou la variance non reconstruite et alors qui permet d'optimiser la reconstruction des variables.

Le modèle du comportement par "ACP" est $\hat{C}$ *où :*

$$\hat{X} = X\hat{C} \quad \tilde{X} = X\tilde{C}$$

Supposons que $C_j^T = \begin{bmatrix} C_{-j}^T & C_{jj}^T & C_{+j}^T \end{bmatrix}$ *est le jeme colonne de* $\hat{C}$
Et v_j *est la variable à reconstruire, où la reconstruction de* v_j *notée* z_j

$$z_j = \frac{1}{1 - C_{jj}} \begin{bmatrix} C_{-j}^T & 0 & C_{+j}^T \end{bmatrix} x(k) \tag{3.15}$$

C_{-j}^T *:les* $(j-1)$ *premiers éléments de la jeme colonne de* $\hat{C}$ C_{+j}^T *:les* $(m-j)$ *derniers éléments de la jeme colonne de* $\hat{C}$

$x(k)$ *le vecteur contenant* m *variables observées du système à l'instant* k *:*

$$x(k) = \begin{bmatrix} x_1(k) & x_2(k) & \cdots & x_m(k) \end{bmatrix}^T$$

Et alors le vecteur de mesure x(k) avec la reconstruction de la jème variable est le suivant :

$$x(k) = \begin{bmatrix} x_1(k) & \cdots & x_{j-1}(k) & z_j(k) & x_{j+1}(k) & \cdots & x_m(k) \end{bmatrix}^T$$

Le but dans cette méthode est de minimiser la variance de l'erreur de reconstruction de la jème variable de $x(k)$:

$$\varepsilon = \xi_j^T \left[x(k) - x_j(k) \right] \tag{3.16}$$

Où $\xi_j^T = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}^T$

Le nombre de composantes principales à retenir s'obtient en minimisant par rapport à ℓ le critère suivant : [voir figure (3.4)]

$$VER(\ell) = \sum_{j=1}^{m} \frac{\sigma_j(\ell)}{\xi_j^T \sum \xi_j} \tag{3.17}$$

avec $\ell = 1 - 1$

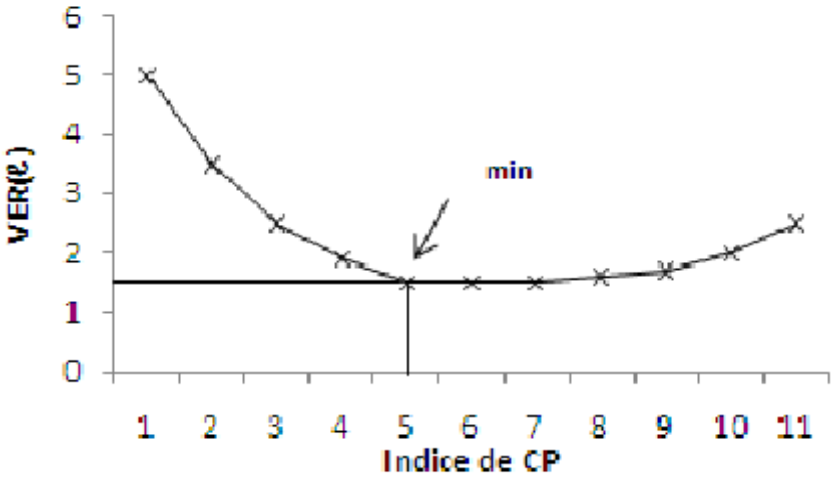


FIGURE *3.4 – Selection du nembre de CPs par VER.*

3.1.3.2 Pourcentage cumulé de la variance (PCV) :

Le principe de cette technique de sélection basé sur l'estimation du pourcentage de la variance de composantes principales. Pour choisir la dimension réduite,il faut

choisir le pourcentage de la variance totale qu'on veut conserver [voir figure (3.5)]

Généralement :

$$PCV(\ell) = 100\left(\frac{\sum_{j=1}^{\ell}\lambda_j}{\sum_{j=1}^{m}\lambda_j}\right)\% \tag{3.18}$$

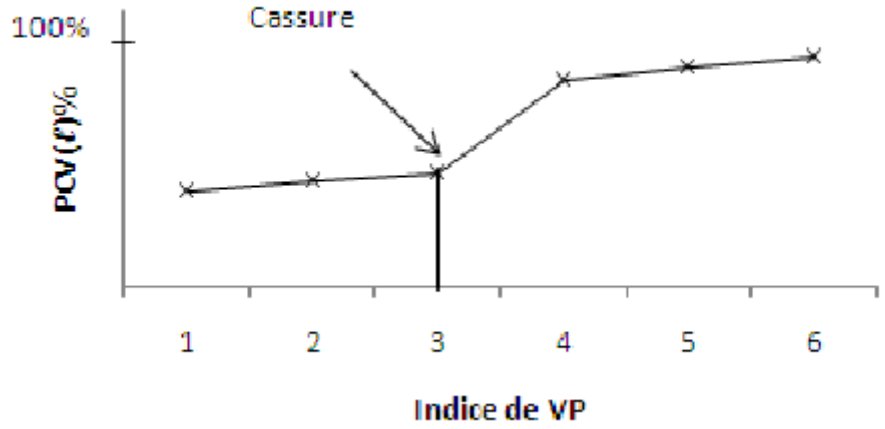


FIGURE *3.5 – Selection du nembre de CPs par PCV.*

La cassure de la courbe entre la 3ème et la 4ème valeur propre, exprime la grande différence entre elles dans la concentration d'information. La figure (3.6) résume la démarche à suivre pour fabriquer un modèle d'un système en fonctionnement normale, à l'aide de la technique de l'analyse en composantes principales linéaire.

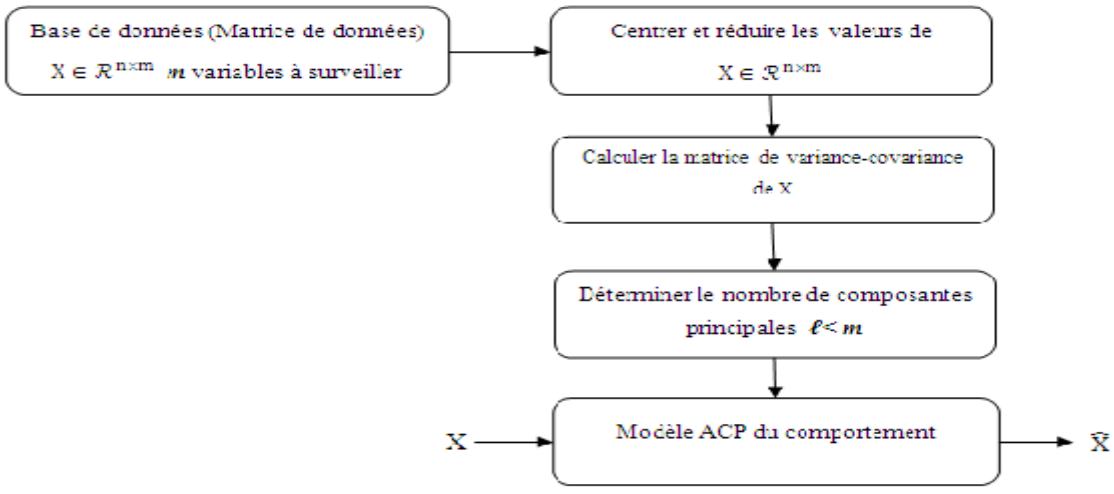


FIGURE *3.6 – Etapes pour la détermination d'un modèle ACP.*

Dans ce qui suit, et après la définition de la structure d'un modèle ACP à partir de données réelles saines, on va voir comment détecter et localiser les défauts qui peuvent apparaître dans le fonctionnement d'un processus, à l'aide du modèle ACP linéaire obtenu.

3.1.4 Détection de nouveauté par ACP Linéaire

Classiquement, plusieurs indicateurs de détection sont utilisés pour la détection d'un fonctionnement anormal par ACP. La plus part des études récentes utilisent l'erreur quadratique de prédiction "Q ", et la statistique de Hotelling "T2", pour la détection des défauts de mesures, ces deux indicateurs donne l'ordre sur l'apparition ou non d'un défaut :

3.1.4.1 Statistique de Hotelling (T^2) *:*

La statistique T^2 *est calculée à partir des* ℓ *premières composantes principales :*

$$T^2(k) = \hat{t}(k)\hat{\Lambda}^{-1}\hat{t}(k) = \sum\nolimits_{a=1}^{\ell} \frac{t_a^2}{\lambda_a} \tag{3.19}$$

Le processus est considéré en fonctionnement anormal (présence d'un défaut) à l'instant k *si :*

$$T^2(k) \geq T^2_{\lim}(k)$$

Où $T^2_{\lim}(k)$ *est la limite du contrôle pour ce critère, L'expression mathématique du seuil pour le critère de Hotelling donnée par :*

$$T^2_{\lim}(k) = \frac{\ell(n+1)(n-1)}{n(n-\ell)} F^2_{\ell,(n-\ell),\alpha} \tag{3.20}$$

$T^2_{\lim}$ *peut être approximée par F-distribution avec* ℓ *le degré de liberté et* α *le niveau de signification (représente le quantile de la distribution).*

3.1.4.2 Statistique SPE (Q) :

Par contre à la première, cette statique est appliquée sur les dernières composantes principales :

$$SPE(k) = \tilde{t}^T(k)\tilde{t}(k) = \sum_{a=\ell+1}^{m} t_a^2 \tag{3.21}$$

Le processus est considéré en fonctionnement anormal (présence d'un défaut) à l'instant k si :

$$SPE(k) \geq \delta^2$$

Où δ^2 *est le seuil de confiance pour cet indicateur.*

$$\delta_\alpha^2 = g\chi_{h,\alpha}^2 \tag{3.22}$$

$$g = \frac{\theta_2}{\theta_2} \tag{3.23}$$

$$h = \frac{\theta_1^2}{\theta_2} \tag{3.24}$$

$$\theta_i = \left(\sum_{i=\ell+1}^{m} \lambda_j^i \right)_{i=1,2} \tag{3.25}$$

δ^2 *peut être approximée par* $\chi^2 - distribution$ *avec h le degré de liberté,* α *le niveau de signification et le coefficient g.*

Pour améliorer la qualité de la détection et réduire les fausses alarmes, un filtre EWMA (Exponentially Weighted Moving Average) peut être appliqué aux résidus :

$$\begin{aligned} &\bar{\tilde{t}}(k) = (I - \beta)\bar{\tilde{t}}(k-1) + \beta\bar{\tilde{t}}(k) \\ &\bar{\tilde{t}}(0) = 0 \end{aligned} \tag{3.26}$$

Où $\beta = \gamma I$ *est une matrice diagonale dont les facteurs d'oubli pour les résidus où* γ *est le facteur d'oubli* $0 < \gamma < 1$. *On obtient ainsi le SPE filtrés :*

$$\overline{SPE}(k) = \bar{\tilde{t}}^T(k)\bar{\tilde{t}}(k) \tag{3.27}$$

Dans ce cas, le processus est considéré en fonctionnement anormal (présence d'un défaut) à l'instant k si : $\overline{SPE}(k) > \bar{\delta}^2$ *où* $\bar{\delta}^2$ *est le seuil de confiance filtré pour cet indicateur filtré qui est donné par la formule suivante :*

$$\bar{\delta}^2 = \frac{\gamma}{2-\gamma}\delta^2 \tag{3.28}$$

3.1.5 Localisation par ACP linéaire

Plusieurs méthodes utilisent l'ACP ont été proposées pour localiser les variables en défaut :

3.1.6 Localisation par calcul de contribution :

Une approche largement utilisée consiste à calculer les contributions des variables à l'indice de détection, et la variable ayant la plus forte contribution à l'instant de détection est la variable affectée. La contribution au critère SPE :

$$cont_j^{SPE}(k) = [\tilde{x}_j(k)]^2 = [x_j(k) - \hat{x}_j(k)]^2 \tag{3.29}$$

$x_j(k)$ *est le jème élément du vecteur de mesures* x *à l'instant* k *donnée. La contribution au critère de Hotelling* (T^2) *:*

$$T^2 = \hat{t}^T \hat{\Lambda}^{-1} \hat{t} = \sum_{a=1}^{\ell} \frac{t_a^2}{\lambda_a} \tag{3.30}$$

$$\lambda_a = S_a^2$$

Nous remarquons que la distance de Hotelling est la somme des carrés des CPs normalisées dont chaque CP ou score est exprimé comme suit :

$$t_a = p_a^T x = \sum_{j=1}^{m} p_{aj} x_j \tag{3.31}$$

p_{aj} *est le* j^{me} *élément du vecteur propre* p_a *correspondent à la valeur propre* λ_a

Et alors, la contribution de la jème variable x_j à une composante principale normalisée est :

$$cont_j^{\left(t_{a/s_a}\right)^2} = \frac{t_a}{\lambda_a} p_{aj} x_j \tag{3.32}$$

Cela prouve que la contribution totale de la variable x_j dans le calcul de toutes les composantes principales normalisées est :

$$cont_j^{T^2} = \sum_{a=1}^{\ell} cont_j^{\left(t_{a/s_a}\right)^2} \tag{3.33}$$

3.1.6.1 Indice de validité des capteurs (SVI) :

Après la détection de la présence d'un défaut, on effectue la reconstruction de toutes les variables à partir de l'instant de détection, et on calcule l'indice de détection. Par exemple si on détecte un défaut à l'instant de temps k, $SPE_j(k)$ représente l'indice de détection de la jème variable reconstruite.

La localisation est effectuée par la comparaison de l'indice de détection avant et après la reconstruction. Cette comparaison est appelée indice de validité $\eta_j(k)$ où :

$$\eta_j(k) = \frac{SPE_j(k)}{SPE(k)} \tag{3.34}$$

Cet indice est toujours inférieur ou égale à 1 $\left(\eta \in \left[\begin{array}{cc} 0 & 1 \end{array}\right]\right)$ parce que l'indice de détection $SPE(k)$ est toujours supérieur ou égale à $SPE_j(k)$. Si le jème capteur est défaillant, l'indice de validité de ce capteur doit être descendant vers le zéro.

Notons que l'équivalence filtrées de ce critère donné par :

$$\bar{\eta}_j^2(k) = \frac{\overline{SPE}_j(k)}{\overline{SPE}(k)} < 1 \tag{3.35}$$

$\Rightarrow$ *la j^{me} variable est affectée par un défaut.*

3.2 ACP à noyau :

Bien que le Acp traite des données linéaires, ACP à noyau s'intéresse aux composants principaux, ou des fonctionnalités, qui sont non linéairement liées aux variables d'entrée. Pour ce faire, nous calculons les produits scalaires dans l'espace des entités à l'aide d'une fonction à noyau dans l'espace d'entrée donnée .tout algorithme qui peut être exprimé par des produits scalaires, c'est-à-dire sans utilisation explicite des variables, la méthode à noyau nous permet de construire une version non linéaire de celle-ci.

Le ACP tente de trouver un sous-espace linéaire à faible dimension auquel les données sont définies. cependant, parfois, les données sont associées à un sous-espace non linéaire à faible dimension où ACP à noyau aura lieu . En regardant la figure (3.7), les données sont principalement situées le long de (ou au moins proche) d'une courbe en 2-D mais le ACP ne peut pas réduire la dimensionnalité de deux à un parce que les points ne sont pas situés le long d'une ligne droite. le ACP à noyau peut reconnaître que ces données sont unidimensionnelles mais dans un espace dimensionnel supérieur, appelé espace caractéristique.le ACP à noyau ne calcule pas explicitement l'espace dimensionnel supérieur, il projéte plutôt les données dans cet espace caractéristique afin que nous puissions classer les données.

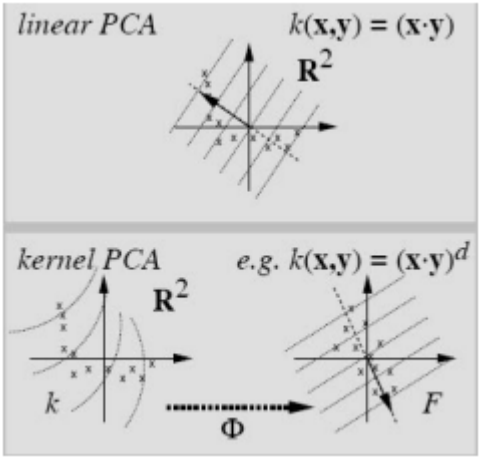


FIGURE *3.7 – ACP à noyau*

3.2.1 Definitions :

Avant d'aller à la détection de la nouveauté par le ACP à noyau, voici quelques termes importants et leurs définitions :

Espace caractéristique :

L'espace caractéristique est un espace étendu par Φ. La méthode du à noyau est utilisée pour sélectionner, à partir des données, un sous-ensemble pertinent formant une base dans un espace caractéristique F. Ainsi, les vecteurs sélectionnés défi nient un sous-espace dans F. En d'autres termes, ils mappent les données de l'espace d'entrée original dans un (en général à haute dimension), espace caractéristique où les relations linéaires existent entre les données.

Matrice Semi-définie positive :

Une matrice Semi-définie positive est une matrice symétrique $n \times n$ A telle que ses valeurs propres sont tous non négatives. Cela s'applique si et seulement si :

$$v^T Av \geq 0$$

pour tous les vecteurs v

Fonction à noyau :

Le noyau est une fonction k pour tous $x, z \in X$ satisfait :

$$k(x, z) = \langle \Phi(x), \Phi(z) \rangle,$$

où Φ un mappage de X à un espace caractéristique (produit scalaire) F

$$\Phi :\rightarrow \Phi(x) \in F.$$

3.2.2 Détection par ACP à noyau :

dans cette partie nous nous intéressons uniquement à la détection de la nouveauté par la méthode à noyau de l'ACP mais avant d'appliquer ceci, nous devons formuler le ACP standard avec des produits scalaires que nous utiliserons alors lors de la dérivation du ACP à noyau , et donc cette formulation consiste à calculer les principaux composants dans F , tout d'abord en calculant la matrice K dans :

$$K_{\iota j} := (\Phi(x_\iota).(\Phi x_j)) \tag{3.36}$$

puis calculer ses vecteurs propres et les normaliser dans F. Ensuite, nous calculons les projections d'un point de test sur les vecteurs :

$$(v^k.\Phi(x)) = \sum_{\iota=1}^{p} \alpha_\iota^k(\Phi(x_\iota).\Phi(x)) \tag{3.37}$$

Notation :

Nous supposons que les données originales sont données en n points, $x_\iota \in \mathbb{R}^d$. Ensuite, il y a des maps qui vont à l'espace caractéristique, F :

$$x_\iota \to \phi(x_\iota) \in F.$$

Si le contexte est clair, nous utiliserons $\phi(x_\iota) = \phi_\iota$. Si x est un point général, alors $\phi(x) = \phi_x$ sera utilisé. De même, nous utiliserons la notation suivante pour la moyenne,la moyenne des points de données extraites :

$$\phi_0 = \frac{1}{n}\sum_{\iota=0}^{n} \phi_\iota \quad et \quad \bar{\phi}_\iota = \phi_\iota - \phi_0.$$

en utilisant l'espace de produit scalaire approprié, nous définissons la fonction à noyau K :

$$k(x, y) = \langle \phi_x, \phi_y \rangle$$

avec la norme dans l'espace caractéristique défini avec le produit scalaire comme

suit :

$$\|\phi_x\|^2 = \langle \phi_x, \phi_x \rangle = k(x, x).$$

comme indiqué précédemment, le produit scalaire dans l'espace caractéristiue peut être calculé directement dans $\mathbb{R}^d$ *en utilisant* K*. Nous définissons la matrice à noyau en utilisant nos* n *points dans* $\mathbb{R}^d$ *et la fonction à noyau* k *:*

$$K_{\iota j} = k(\mathbf{x}_{\iota,}\mathbf{x}_j)$$

afin que nous puissions calculer la moyenne de la matrice à noyau extraite en utilisant le calcul précédent pour la moyenne du produit scalaire extraité .C'est :

$$\tilde{K}_{\iota j} = k\langle \tilde{\phi}_\iota, \tilde{\phi}_j \rangle$$

$$= k(\mathbf{x}_\iota, \mathbf{x}_j) - \frac{1}{n}\sum_{r=1}^{n} k(\mathbf{x}_\iota, \mathbf{x}_r) - \frac{1}{n}\sum_{s=1}^{n} k(\mathbf{x}_j, \mathbf{x}_s) + \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} k(\mathbf{x}_r, \mathbf{x}_s).$$

ces valeurs peuvent donc être calculées en utilisant la matrice à noyau originale K*.*

Les valeurs propres et les vecteurs propres de la matrice de covariance :

Pour effectuer l'analyse des composants principaux dans l'espace caractéristique F*, on définie la matrice de covariance , comme avant, dans l'espace caractéristique comme suit :*

$$C = \frac{1}{n}\sum \tilde{\phi}_\iota \tilde{\phi}_\iota^T = \frac{1}{n}\Phi\Phi^T$$

où Φ *est une matrice dont les colonnes sont données par la moyenne des données extraites dans l'espace caractéristique :*

$$\Phi = [\tilde{\phi}_1, \phi_2, ..., \tilde{\phi}_n].$$

en utilisant cette notation, nous pouvons écrire :

$$\tilde{K} = \Phi^T \Phi.$$

Les données dans l'espace caractéristique sont généralement de très grande dimension. L'espace caractréristiqueé est d'une dimension infinie lors de l'utilisation du noyau Gaussien (Gaussian kernel) ci-dessous, donc nous ne voulons pas construire les vecteurs propres réels $\mathbf{v}$*, au lieu de cela, nous calculerons seulement le produit linéaire impliquant* $\mathbf{v}$ *:*

Maintenant, supposons que $\mathbf{v}$ *est le vecteur propre de la matrice de covariance* $\mathbf{C}$ *et* λ *sa valeur propre associée (*λ *et* $\mathbf{v}$ *s'appellent eigenpairs), alors :*

$$\lambda v = Cv.$$

si nous développons $Cv = \frac{1}{n}\Phi(\Phi^T v)$*, nous voyons que* $\mathbf{v}$ *est une combinaison linéaire des colonnes de* Φ*, donc* $\mathbf{v}$ *se trouve dans les coordonnées* $S = \{\tilde{\phi}_1, \tilde{\phi}_2, ..., \tilde{\phi}_n\}$ *ainsi, nous pouvons considérer l'équivalent système d'équations :*

$$\lambda \Phi^T v = \Phi^T C v. \tag{3.38}$$

nous définissons α *les coordonnées de* $\mathbf{v}$ *par rapport aux vecteurs caractéristiques :*

$$v = \sum_{\iota=1}^{n} \alpha_\iota \tilde{\phi}_\iota = \Phi\alpha. \tag{3.39}$$

par la définition de la covariance C*, et en remplaçant* (3.39) *par* (3.38)*, on a :*

$$\lambda \Phi^T \Phi\alpha = \frac{1}{n}\Phi^T \Phi \Phi^T \Phi\alpha.$$

de là, nous obtenons notre équation principale :

$$n\lambda \tilde{K}\alpha = \tilde{K}^2\alpha. \tag{3.40}$$

si nous comparons les solutions à celles ci-dessus avec les solutions à la valeur

propre suivante , probléme du vecteur propre :

$$n\lambda\alpha = \tilde{K}\alpha, \tag{3.41}$$

Nous voyons que toute solution à (3.41) *sera clairement une solution à* (3.40). *À l'inverse, nous pouvons avoir des solutions à* (3.40) *qui ne sont pas des solutions à* (3.41), *mais ces vecteurs seraient nécessairement dans l'espace nul de* $\tilde{K}$, *Ainsi, nous avons donc notre résultat principal plus Loin, nous avons rendu* α *et* λ *calculable à partir de la matrice à noyau* $\tilde{K}$.

Mise à l'échelle des vecteurs propres :

Nous devons veiller à ce que nous ayons des vecteurs propres unitaires $\mathbf{v}$, *de sorte que la projection d'un nouveau le vecteur en* F *donnera la formule dans* (3.41). *Si nous utilisons* (3.39) *et la définition de* $\tilde{K}$, *puis :*

$$|v\|^2 = \langle \Phi\alpha, \Phi\alpha \rangle = \alpha^T \Phi^T \Phi\alpha = \alpha^T \tilde{K}\alpha = n\lambda\alpha^T\alpha = n\lambda \|\alpha\|^2 .$$

par conséquent, nous allons mettre à l'échelle α *pour que :*

$$\|\alpha\| = \frac{1}{\sqrt{n\lambda}}.$$

Variance résiduelle :

Nous pouvons utiliser les valeurs propres de la matrice de covariance pour déterminer quel pourcentage du total la variation est expliquée en prenant $\mathbf{k}$ *hors de* $\mathbf{n}$ *vecteurs propres possibles. Ceci est donné par la somme des premières valeurs propres normalisées de* K *:*

$$\sum_{\iota=1}^{k} \frac{\lambda_\iota}{\sum_{j=1}^{n} \lambda_j}.$$

Une façon de calculer ceci consiste à utiliser la trace d'une matrice (la somme de ses éléments diagonaux). pour une $\mathbf{n} \times \mathbf{n}$ *matrice* A *avec des valeurs propres*

réelles,

$$\sum_{j=1}^{n} \lambda_j = tr(A) = \sum_{j=1}^{n} A_{jj.}$$

En particulier, si les vecteurs propres proviennent de la matrice symétrique $\tilde{K}$ et nous utilisons p eigenpairs pour notre sous-espace approximatif, alors la variance résiduelle est le pourcentage de variation expliqué en utilisant p dimensions dans l'espace caractéristique.

Mesure de la nouveauté d'un nouveau point :

Compte tenu d'un nouveau vecteur $\mathbf{z} \in \mathbb{R}^d$*, nous définissons la nouveauté comme une erreur de reconstruction légèrement modifiée dans l'espace caractéristique :*

$$F(z) = \left\|\bar{\phi}_z\right\|^2 - \left\|Proj_Q(\bar{\phi}_z)\right\|^2$$

où Q est un q sous-espace dimensionnel de F dont la base est donnée par les vecteurs propres $v_1, .., v_q$*.*

Nous montrerons comment calculer la valeur de F sans calculer explicitement $\bar{\phi}_z$ *, Ou l'une des vecteurs propres. Tout d'abord, nous regardons à* $\left\|\bar{\phi}_z\right\|^2$ *:*

$$\begin{aligned}\left\|\bar{\phi}_z\right\|^2 &= \langle \phi_z - \phi_0, \phi_z - \phi_0 \rangle \\ &= \langle \phi_z, \phi_z \rangle - 2\langle \phi_z, \phi_0 \rangle + \langle \phi_0, \phi_0 \rangle \\ &= \langle \phi_z, \phi_z \rangle - \tfrac{2}{n} \textstyle\sum_{\iota=1}^{n} \langle \phi_z, \phi_\iota \rangle = \tfrac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} \langle \phi_r, \phi_s \rangle \\ &= k(z,z) - \tfrac{2}{n} \textstyle\sum_{\iota=1}^{n} k(z, x_\iota) = \tfrac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} k(x_r, x_s).\end{aligned}$$

Chaque expression est maintenant calculable simplement en utilisant la fonction à noyau dans $\mathbb{R}^d$*. Aller à la prochaine projection, si nous développons la projection comme :*

$$Proj_Q(\bar{\phi}_z) = \beta_1 v_1 = \beta_2 v_2 + ... + \beta_q v_{q,}$$

Alors depuis que les vecteurs propres sont orthonormés dans l'espace caractéristique,

$$\left\|Proj_Q(\bar{\phi}_z)\right\|^2 = \beta_1^2 + ... + \beta_q^2 = \langle \bar{\phi}_z, v_1 \rangle^2 + \langle \bar{\phi}_z, v_q \rangle^2.$$

Typiquement, l'extraction d'une "feature" à partir du nouveau point de données

$\bar{\phi}_z$, *en utilisant le vecteur propre* $\mathbf{v}$, *Signifie le calcul de la projection scalaire :*

$$\langle \bar{\phi}_z, v \rangle.$$

l'aide de la définition de $\mathbf{v}$ *dans* (3.39), *cette "feature" (ou projection scalaire) peut être calculée sans calculer directement* $\mathbf{v}$. *De* (3.41), *les coordonnées de* $\mathbf{v}$ *dans le vecteur* α *ont été calculé à partir de la matrice* K *à noyau extraite. Ainsi :*

$$\langle \bar{\phi}_z, v \rangle = \langle \bar{\phi}_z, \sum_{\iota=1}^{n} \alpha_\iota \bar{\phi}_\iota \rangle = \sum_{\iota=1}^{n} \alpha_\iota \langle \bar{\phi}_z, \bar{\phi}_\iota \rangle. \tag{3.42}$$

Ce dernier produit scalaire a déjà été calculé comme suit :

$$\langle \bar{\phi}_z, \phi_\iota \rangle = k(z, x_\iota) - \frac{1}{n}\sum_{r=1}^{n} k(x_\iota, x_r) - \frac{1}{n}\sum_{s=1}^{n} k(z, x_s) + \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} k(x_r, x_s). \tag{3.43}$$

Détection de nouveauté :

Pour chaque nouveau point z, *on peut calculer la "nouveauté",* $F(z)$, *en regardant la valeur maximale de* F *sur les valeurs* z. *De là, nous voyons quelles valeurs* z *sont "normal "pour obtenir une valeur du seuil* v. *Une valeur* z *est considérée comme nouvelle si :*

$$F(z) \geq v.$$

3.3 Conclusion

Nous avons présenté le principe de l'ACP linéaire et le diagnostic basé sur un modèle ACP qui servent pour la détection de nouveauté . Deux indices de détection des anomalies ou nouveautés de façon globale, T2 et SPE, et deux critères de localisation des défauts, localisation par calcul de contribution et par l'indice de validité des capteurs, sont expliqués dans cette partie. Et concernant la sélection du nombre optimal de CPs à retenir, nous avons détaillé les deux critères les plus efficaces, l'un consiste à étudier le pourcentage cumulé de la variance et l'autre basé sur la minimisation de l'erreur de reconstruction. quant au ACP à noyau qui traite les systèmes non linéaires, qui consiste à trouver la linéarisation par les produits

scalaires et à calculer les principales projections de composants sur les vecteurs propres tout en commençant par le calcul de la matrice à noyau et à l'aide de la fonction à noyau pour enfin passer à la détection de nouveauté .

CHAPITRE 4

CONTRÔLE AVANCÉ DES RÉSEAUX D'APPROVISIONNEMENT EN EAU

4.1 Introduction

La modélisation des processus non linéaires complexes est une tâche difficile en raison des non-linéarités inhérentes, des caractéristiques variables dans le temps. Plusieurs méthodes de modélisation ont été étudiées, en particulier celles basées sur des modèles mathématiques traditionnels en termes de description mécanique du système. Au cours des deux dernières décennies, plusieurs approches incluant des techniques d'intelligence computationnelle ont été appliquées avec succès, parmi ces techniques, la modélisation et le contrôle flou ont été testés sur plusieurs processus en utilisant des données représentatives et qualitatives.

Au cours des dernières années, de nombreux chercheurs ont fait de grands efforts pour la modélisation des Systèmes de distribution d'eau potable (DWDS) pour améliorer ses performances. Divers modèles mathématiques basés sur la masse, les lois sur la conservation de l'énergie et les modèles axés sur les données du DWDS ont été établis. Afin de s'assurer que le DWDS peut fonctionner avec succès, il est nécessaire d'examiner sa quantité et ses performances de qualité. Ainsi, la construction d'un modèle adapté est importante pour l'étude des réseaux de distribution d'eau.

Une opération robuste et sûre des réseaux d'approvisionnement en eau doit être satisfaite à chaque fois, de sorte que les capteurs et les actionneurs doivent être utilisés efficacement. En raison de la dynamique non linéaire décrite dans le modèle de qualité et du problème de contrôle contraint multivariable, le modèle de contrôle prédictif (MPC) est sélectionné pour implémenter le contrôle de qualité à DWDS.

L'article est organisé de la manière suivante. Le modèle mathématique de DWDS est élaboré en section 4.2, Les résultats sont présentés dans la section 4.3. Le ACP est ensuite appliqué sur notre systéme dans la section 4.4 Enfin, la conclusion

est tirée en section .

4.2 Modèle mathématique

Un système de distribution d'eau est l'ouvrage physique qui fournit de l'eau de la source d'eau au point final ou à l'utilisateur prévu. Il est conçu pour fournir une quantité et une qualité d'eau suffisantes pour répondre aux exigences du client. Le système d'alimentation en eau se compose d'un grand nombre de composants inter connectés. Ils comprennent des éléments qui peuvent modifier le débit d'eau dans le système appelé actif (par exemple, les vannes et les pompes) et les éléments passifs (ne modifiez pas le débit comme les tuyaux, les réservoirs et les réservoirs).

La conservation de la masse et de l'énergie sont les lois qui régissent l'écoulement dans les systèmes de conduites en condition stable.

- *Conservation de la masse*

La loi de conservation des états de masse indique que le taux de stockage dans un système est égal à la différence entre l'entrée et la sortie du système. Pour un nœud de jonction.

$$\frac{dS_i\left(t\right)}{dt} = \sum Q_i\left(t\right) - V_i\left(t\right)$$

Où Q_i Est l'afflux total dans le nœud i $[m^3/s]$, V_i Le volume d'eau au nœud i $[m^3/s]$, et dS_i/dt Est le changement de stockage $[m^3/s]$.

- *Conservation d'énergie*

La loi sur la conservation de l'énergie est généralement exprimée en termes de changement de tête le long d'une boucle ou d'un chemin d'énergie. Considérons une section de tuyau avec longueur l_p $[m]$, Section transversale A_p $[m^2]$, et Δh_p La différence de charge entre les deux extrémités du tuyau,g Est l'accélération gravitationnelle, h_loss Représente la perte de Charge via composant i sur le chemin. L'évolution du flux $Q_p(t)$ $[m^3/s]$ À travers le tuyau est donné par :

$$\frac{dQ_p\left(t\right)}{dt} = \frac{gA_p}{l_p}\left(\Delta h_p\left(t\right) - h_{loss}\left(t\right)\right)$$

Les pertes de charge représentent les pertes d'énergie le long de la section de tuyauterie sont décrites comme suit :

$$h_{loss}(t) = h_{loss.f_p}(t) + h_{loss.m}(t)$$

Où $h_{loss.f_p}$ Représente des pertes par frottement et $h_{loss.m}$ Les pertes locales mineures [36]. Pour obtenir un modèle dynamique de DWDS, les paramètres des composants du réseau doivent être connus ou estimés, les équations de masse et d'énergie sont utilisées pour développer une représentation d'espace d'état de la forme suivante :

$$\dot{x}(t) = f(x(t), u(t), d(t))$$

Nous considérons que : x Est l'état (le flux entre les tuyaux et les têtes dans les nœuds), u Est l'entrée de contrôle (coefficient de perte des soupapes et injection de pression des pompes)d Est l'entrée de perturbation exogène (le modèle de consommation), et f Est la fonction de transition d'état non linéaire.

- *Modèle de qualité*

Pour éliminer les microorganismes qui causent des problèmes d'écoulement, le nombre de désinfectants chimiques peut être utilisé dans le DWDS, par exemple, comme il est peu coûteux et facilement appliqué à l'eau, il doit être conservé dans certaines limites. Un modèle de qualité de réseau d'eau est essentiel à l'estimation de la qualité de l'eau. Lorsqu'une substance est injectée dans le DWDS, elle se propage dans la direction du mouvement de l'eau et elle peut réagir et se décomposer dans le temps à un certain rythme. L'équation décrivant le transport et la carie du chlore dans un flux d'eau linéaire est la suivante :

$$\frac{\partial C_i(t,d)}{\partial t} = -v\frac{\partial C_i(t,d)}{\partial d} + R_i C_i(t,d)$$

Où t Est l'instant-temps au cours de la période de marche hydraulique,C_i Est la concentration de chlore dans le tuyau i , d Est la distance du nœud initial,v Signifie l'écoulement de la vitesse dans le tuyaui et R_i Est le coefficient de vitesse de réaction dans le tuyau i (Il est négatif en raison de la désintégration de la

concentration de chlore dans le temps). Habituellement, l'approche lagrangienne est utilisée pour la simulation du transport et la concentration du chlore dans l'eau. Par conséquent, le logiciel peut calculer l'information hydraulique et ensuite résoudre le modèle de contaminant lagrangien en calculant la concentration à chaque lien du réseau en utilisant des mesures réelles :

$$y\left(k\right) = f\left(F; d_c, u\left(k\right)\right)$$

$$y\left(k\right) = f\left(F; d_c, u\left(k\right)\right)$$

Où N Est le nombre de noeuds dans le réseau,$y(k) \in N$ *Est un vecteur des données de concentration de chlore à chaque nœud, la fonction f Est l'algorithme de solution hydraulique et de qualité. Cette fonction prend comme graphique d'arguments F Qui représente les nœuds et les tuyaux du réseau, les exigences du consommateur* $d_c \in N$*, et* $y(k) \in N^I$ *est la concentration de substance injectée aux noeuds* N^I*.*

4.3 Étude de cas

Une référence de deux nœuds d'injection de chlore et de deux nœuds surveillés est étudiée Fig.4.1. Il y a 16 nœuds, 27 tuyaux et 3 réservoirs de stockage. L'eau est pompée de la source (nœud 100 et nœud 200) par deux pompes (pompe 201 et pompe 101) et est également fournie par les réservoirs (nœud 17,18,19). La concentration de chlore au nœud 16 et 8 est les deux sorties y_1 *et* y_2*. La concentration de chlore aux nœuds d'injection 5 et 10 sont les entrées* u_1 *et* u_2*.*

Pour simuler la dynamique hydraulique et dynamique du DWDS, l'horizon de la demande d'eau est réglé à 55 heures avec un intervalle de 15 minutes. Fig. 4.2 et Fig. 4.3 Montre les séries chronologiques des données d'entrée selon une perspective de commande et et des données de sortie.

Les liens sont utilisés sur un VCS afin de créer des connexions de routage entre une sous-zone et une autre sous-zone ou zone. Chaque extrémité du lien est connue sous le nom de nœud. Les canaux sont appliqués aux liens afin de limiter la bande

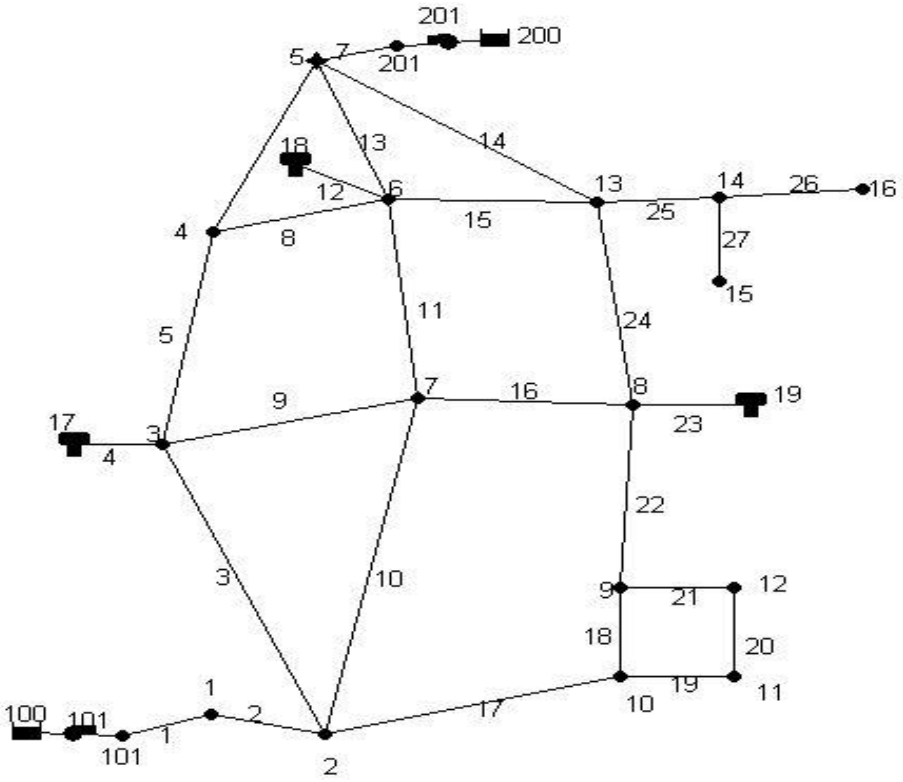


FIGURE *4.1 – Un réseau de distribution d'eau potable DWD*

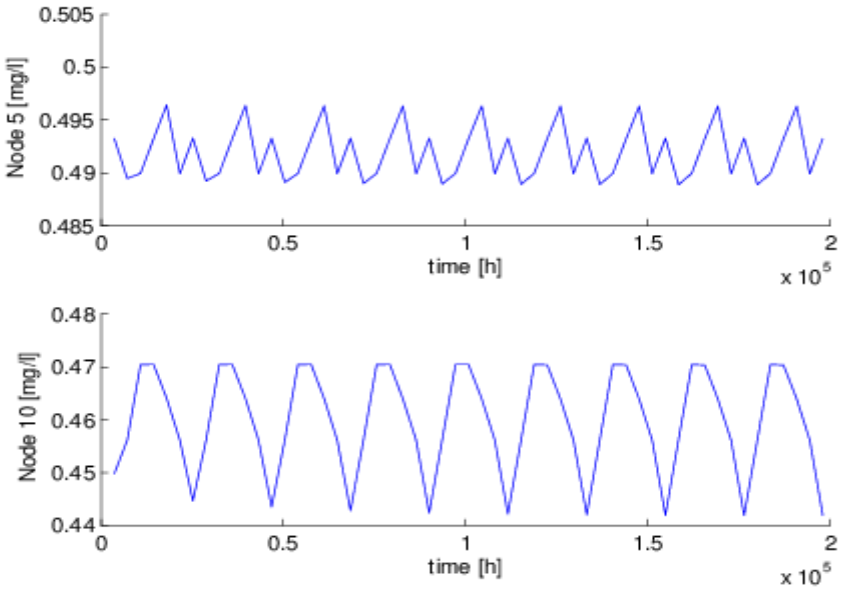
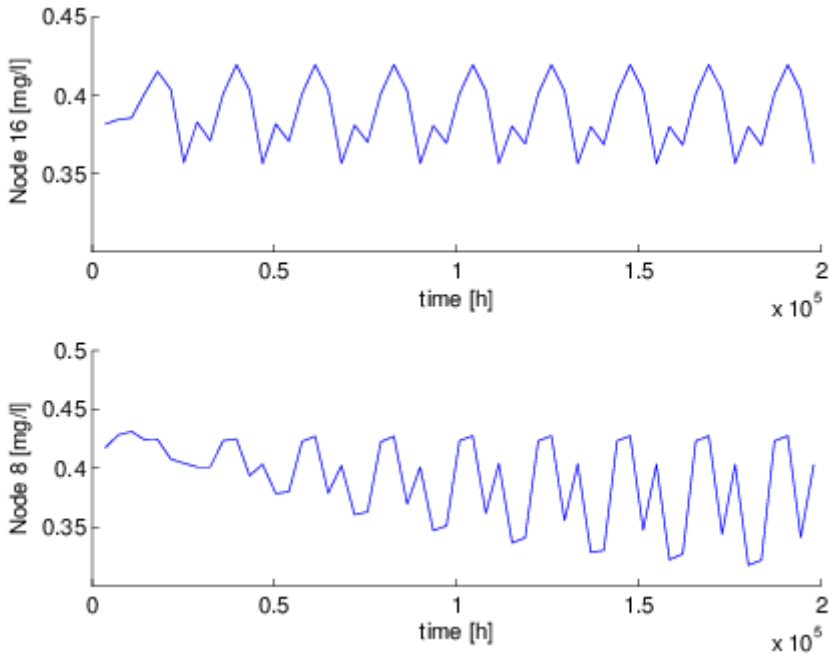


FIGURE *4.2 – Données d'entrée*



passante disponible entre les deux nœuds. *la figure suivante montre l'ensemble des débits en lien* 1, 7, 14, 17 *:*

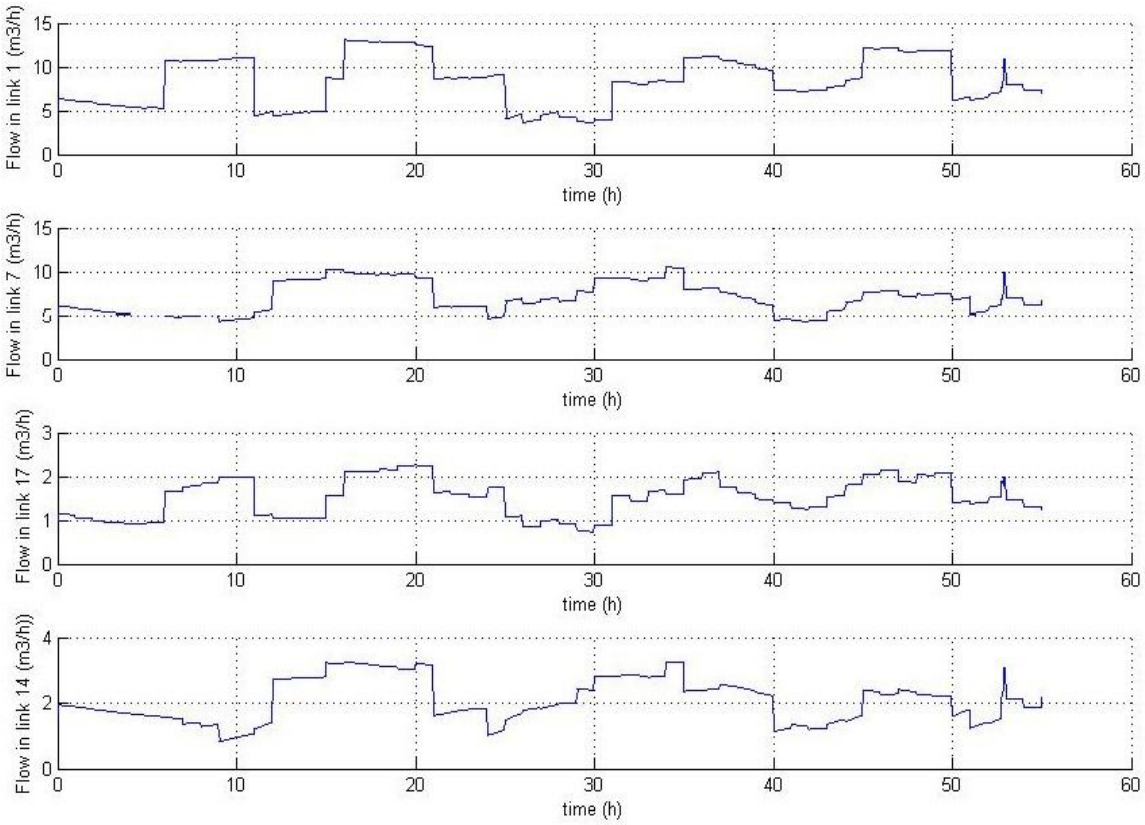


FIGURE *4.4 – Ensemble des débits*

4.4 Tests et résultas

on considére que notre jeu de données obtenu par les capteurs et les actionneurs est composé de 8 variables (4 concentrations de chlore (u1, u2, y1, y2) et 4 débits (q1, q7, q17, q14)) et de 1105 observations à 55 h d'intervalle ,

Phase de la construction du modéle

On applique notre approche ACP sur notre ensemble de données d'apprentissage pour construire le modèle, la figure 4.5 montre les valeurs propres calculées en fonction d'indice des composants principaux :

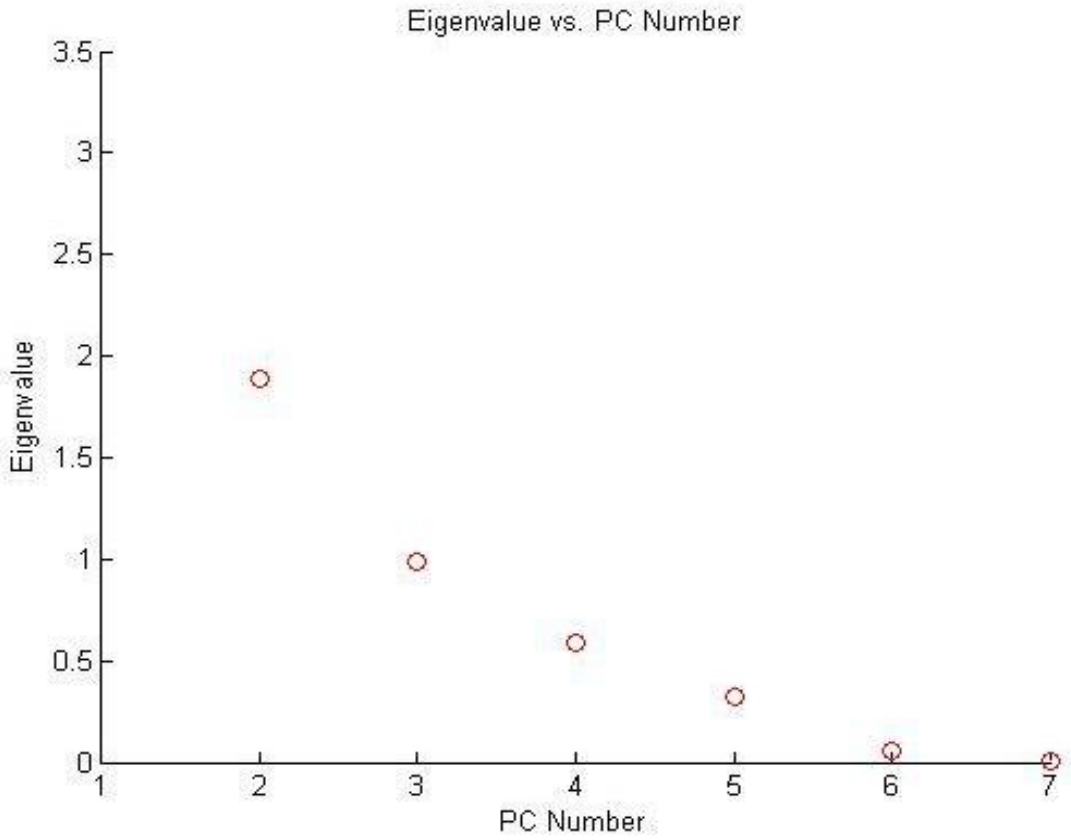


FIGURE *4.5 – Valeurs propores*

Pour ensuite déterminer le nombre de CPs à retenir qui est égal à 2 et qui donne la nouvelle dimension de l'espace réduit par le choix de l'approche du PCV (dans le cas de notre application c'est à 95%) .

Aprés, on utilise la statistique SPE(Q) dans le mode normale (absence du défault) répresenté sur la figure 4.6 avec un seuil égale a 5.3925 :

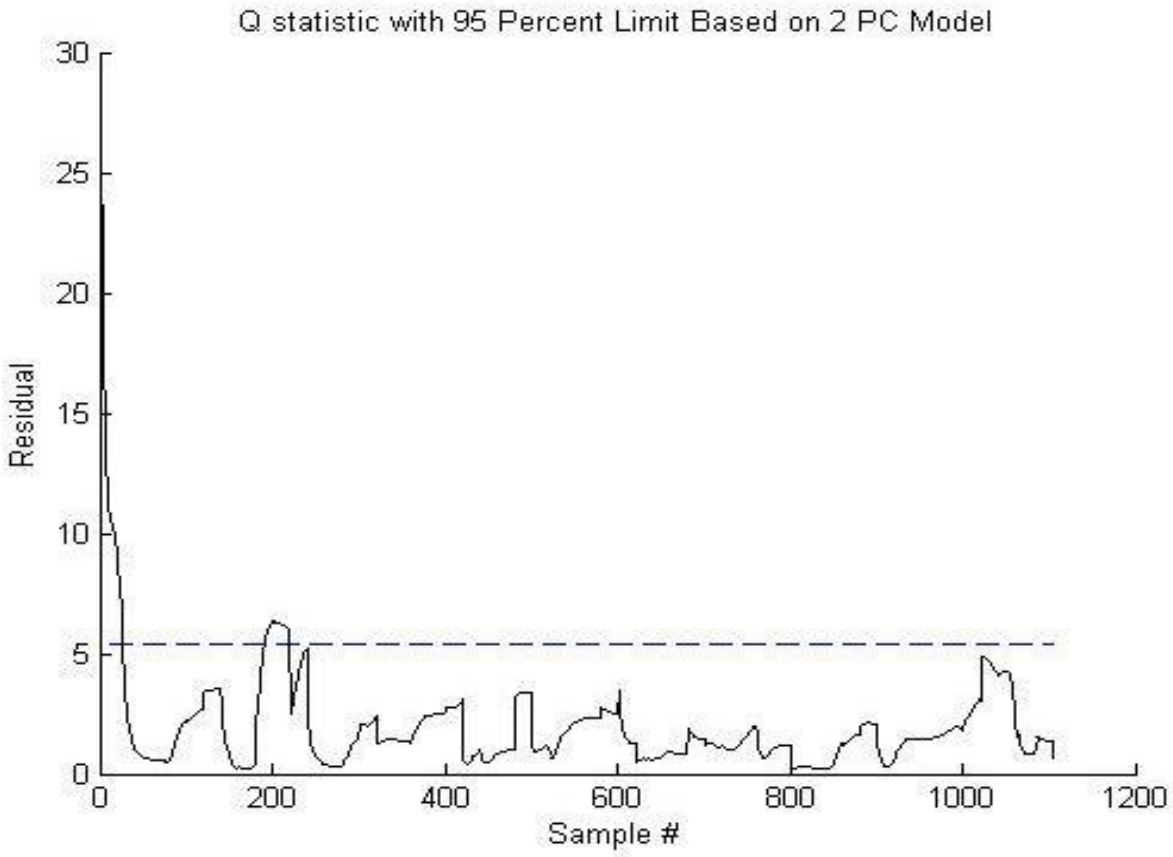


FIGURE *4.6 – la statistique SPE(Q) avec limite à 10% basé sur un modéle de 2 CP*

Phase de tests et défauts

Avant de créer un défaut sur un capteur on applique la statistique d'Hotelling (T^2) sur notre donnée de test qui se constitue de 525 observations dans le mode normal et qui est représenté dans la figure 4.7 avec un seuil T^2_{lim} égale à 6.057 :

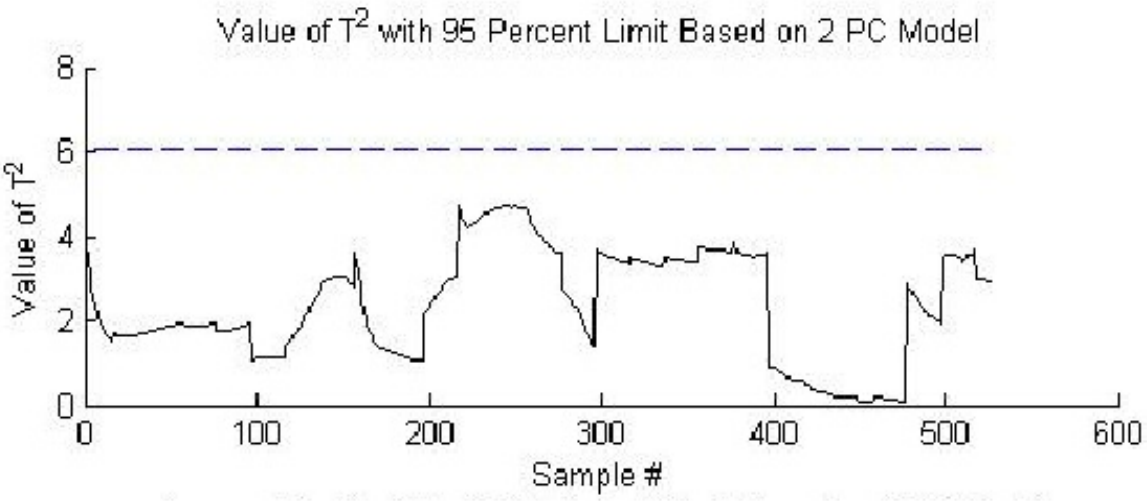


FIGURE *4.7 – la statistique de Hotelling (T^2) avec limite à 10% basé sur un modéle de 2 CP*

On tente de créer un défaut sur le deuxième capteur au 450ème échantillon qui est dans notre cas le capteur de la concentration du chlore injecté au niveau du nœud 10 (entrée de commande u2) , avec notre méthode on est de mesure de détecter ce comportement anormale par le dépacement du seuil à cet instant comme on le voit dans la figure 4.8 ,

Et avec la localisation par le calcul de contribution on peut localiser le défaut qui se trouve sur le capteur, dans la figure 4.8 on remarque que la variable 2 prend la plus grande amplitude ce qui vient à dire que ce défaut se trouve sur ce dernier .

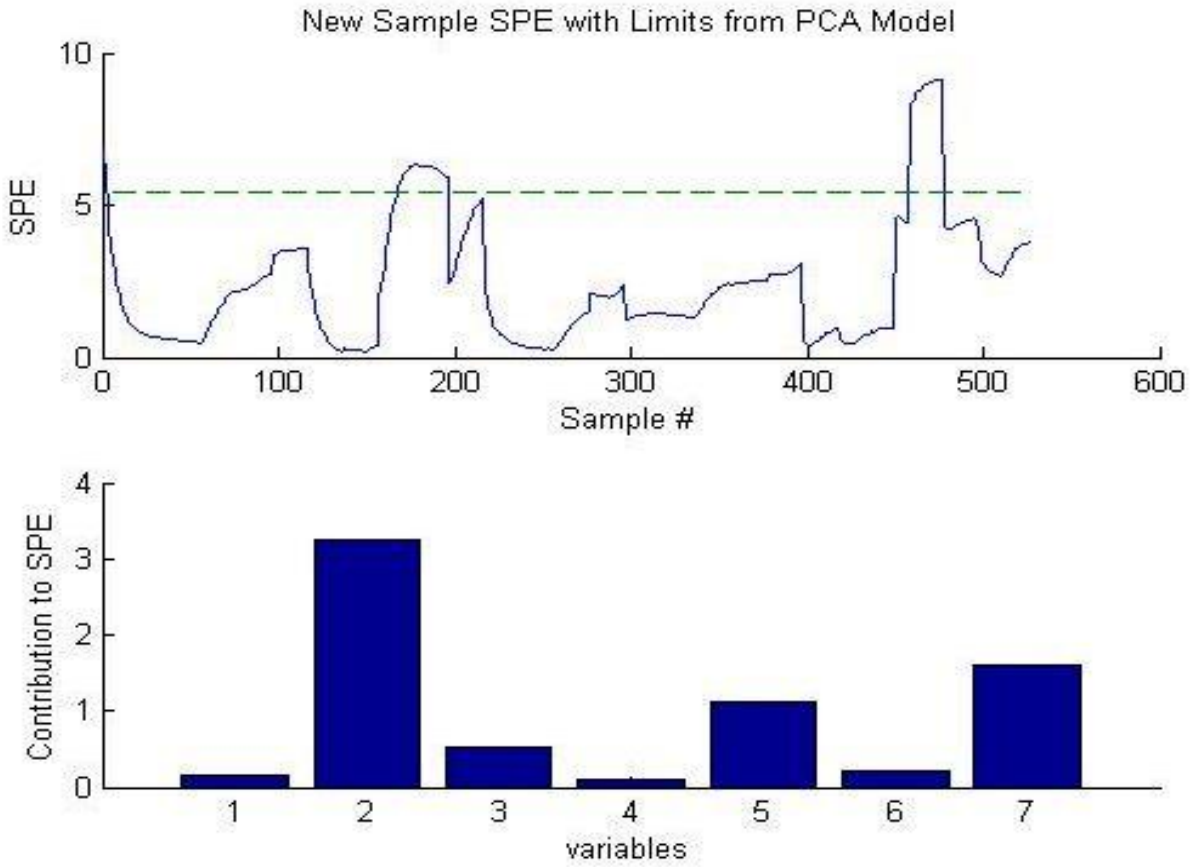


FIGURE *4.8 – Détection et localisation du défaut sur le capteur 2*

On peut même déduire de cette figure que ce défaut est du type fugitif car il apparaît a l'instant désiré et disparaît subitement .

4.5 Conclusion

Cet article montre que le contrôle de supervision peut être appliqué avec succès pour maintenir la concentration de chlore dans les réseaux d'eau potable dans un intervalle prédéfini. Les variables manipulées sont calculées afin d'optimiser le comportement futur des processus sous-jacents, concernant la précision du point de consigne et la minimisation de l'énergie. Où N Est le nombre de nœuds dans le réseau,$y(k) \in N$ Est un vecteur des données de concentration de chlore à chaque nœud, la fonction f Est l'algorithme de solution hydraulique et de qualité. Cette fonction prend comme graphique d'arguments F Qui représente les nœuds et les

tuyaux du réseau, les exigences du consommateur $d_c \in N$, et $y(k) \in N^I$ est la concentration de substance injectée aux nœuds N^I.

Les résultats et la discussion de l'application de l'analyse en composantes principales sur notre système de distribution d'eau potable afin de détecter et localiser les défauts de capteurs, ont montré le succès et l'efficacité de cette approche.

CONCLUSION GÉNÉRALE

Dans ce mémoire, nous avons passer en revue les différentes approches les plus utilisées pour sécuriser les systèmes SCADA à cause de leur vulnérabilité contre les attaques affectantes au niveau du protocole de communication, et particulièrement la détection d'intrusion par anomalie qui est forcément une nouveauté, cette dernière comporte plusieurs techniques et méthodes de classification par des algorithmes afin de construire un modèle de décision en passant par la sélection des attributs et la construction des classifieurs, et donc en montrant aussi que la détection d'intrusion dans unsystème de contrôle peut être vue comme un problème de classification.

L'ACP est l'une des approches de détection de nouveauté à base de données, et comme on peut le voir, c'est une technique statistique descriptive dont le principe est simple mais qui met en œuvre des calculs numériques importants, pour cette raison elle n'a pu se développer qu'avec l'apparition des ordinateurs, elle est utilisé pour détecter et localiser les défauts de capteur.

Généralement le principe de cette technique est la réduction de la dimensionnalité d'une base de données avec la perte du moins d'information possible, pour faciliter l'analyse, cette approche permet d'identifier les relations entre les variables du système afin de trouver un modèle prévu de bon fonctionnement aide par la suite à détecter et localiser les défauts de capteurs par génération des indicateurs de défauts (résidu), en comparant le comportement donné par les variables mesurées et le comportement donné par le modèle ACP prévu. Deux indicateurs des anomalies qui sont certainement des nouveautés, l'indicateur d'Hotteling T^2 et la statique de l'erreur quadratique SPE, sont décrits pour la détection, où le premier indicateur est calculé par les premières composantes principales et le deuxième est calculé par les dernières composantes qui représentent les résidus.

Concernant la localisation où l'isolation des capteurs défaillants, nous avons présenté les méthodes les plus importantes parmi un nombre important de méthodes de localisation, et sont respectivement : la localisation par calcul de contribution à l'indice de détection et la localisation par l'indice de validité des capteurs.

Quant aux systèmes non linéaires, l'ACP à noyau qui s'intéresse aux composants principaux, qui sont non linéairement liés aux variables d'entrée permet de construire une version non linéaire de celle-ci en formulant l'ACP standard avec des produits scalaires tout en commençant par le calcul de la matrice à noyau et à l'aide de la fonction à noyau pour ensuite calculer les principales projections de composants sur les vecteurs propres pour enfin s'en servir à la détection de nouveauté .

Dans notre travail qui consiste à maintenir la concentration de chlore dans notre réseau de distribution d'eau potable pour un bon approvisionnement en eau, on a utilisé l'ACP pour détecter et localiser les défauts sur les capteurs par notre approche de détection de nouveauté, Les résultats montrent que le contrôle de supervision peut être appliqué avec succès, les variables manipulées ont été calculé pour optimiser le comportement futur des processus sous-jacents, en ce qui concerne la précision du point de consigne et la minimisation de l'énergie. ainsi le succès et l'efficacité de cette méthode (ACP) qui est récemment plus utilisée .

BIBLIOGRAPHIE

[1] D. E. Denning. An intrusion-detection model. IEEE Transactions on Software Engineering, SE-13(2) :222-232, Feb 1987.

[2] A. Valdes and S. Cheung. Communication pattern anomaly detection in process control systems. In IEEE Conference on Technologies for Homeland Security,2009. HST'09., pages 22-29. IEEE, 2009.

[3] O. Linda, T. Vollmer, and M. Manic. Neural network based intrusion detection system for critical infrastructures. In International Joint Conference on Neural Networks, 2009. IJCNN 2009., pages 1827-1834, June 2009.

[4] N. Fovino, A. Carcano, T. D. L. Murel, A. Trombetta, and M. Masera. Modbus/dnp3 state-based intrusion detection system. In Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), pages 729-736, April 2010a.

[5] D. Wei, Y. Lu, M. Jafari, P. M. Skare, and K. Rohde. Protecting smart grid automation systems against cyberattacks. IEEE Transactions on Smart Grid, (99) :1-1, 2011.

[6] C. Alcaraz and J. Lopez. Wasam : A dynamic wide-area situational awareness model for critical domains in smart grids. Future Generation Computer Systems,30 :146-154, 2014a.

[7] P. Gross, J. Parekh, , and G. Kaiser. Secure selecticast for collaborative intrusion detection systems. In Proceedings of the 3rd International Workshop on Distributed Event-Based Systems (DEBS), page 50, 2004.

[8] [4] Hodge, V, J, et Austin, J. (2004), A survey of outlier detection methodologies, Artificial Intelligence Review, 22(2), 85-126.

[9] Agyemang, M., Barker, K., and Alhajj, R. 2006. A comprehensive survey of numeric and symbolic outlier mining techniques. Intelligent Data Analysis 10, 6, 521-538.

[10] Markou, M, Singh, S, December 2003a. Novelty detection : a review. part 1 : statistical approaches. Signal Processing 83 (12), 2481-2497.

[11] Markou, M, Singh, S, December 2003b. Novelty detection : a review. part 2 : neural network based approaches. Signal Processing 83 (12), 2499-2521.

[12] Patcha, A. et J-M. Park (2007). An overview of anomaly detection techniques : Existing solutions and latest technological trends. Comput. Networks 51, 3448-3470.

[13] Beckman, R. J. and Cook, R. D. 1983. Outlier...s. Technometrics 25, 2, 119-149.

[14] Abubakar, M. ; Doma, U. D. ; Kalla, D. J. U. ; Ngele, M. B. ; Augustine, C. L. D., 2006. Effects of dietary replacement of maize with malted or unmalted sorghum on the performance of weaner rabbits. Livest. Res. Rural Dev., 19 (5) : 65. 6pp.

[15] Rousseeuw, P. J. and Leroy, A. M. 1987. Robust regression and outlier detection. John Wiley and Sons, Inc., New York, NY, USA.

[16] Barnett, V. and Lewis, T. 1994. Outliers in statistical data. John Wiley and sons.

[17] awkins, D. 1980. Identiffication of outliers. Monographs on Applied Probability and Statistics. Bakar, Z., Mohemad, R., Ahmad, A., and Deris, M. 2006. A comparative study for outlier detection techniques in data mining. Cybernetics and Intelligent Systems, 2006 IEEE Conference on , 1.6.

[18] ong, X., Wu, M., Jermaine, C., and Ranka, S. 2007. Conditional anomaly detection. IEEE Transactions on Knowledge and Data Engineering 19,5, 631-645.

[19] Salvador, S. and Chan, P. 2003. Learning states and rules for time-series anomaly detection.Tech. Rep. CS-2003-05, Department of Computer Science, Florida Institute of Technology Melbourne FL 32901. march.

[20] Shekhar, S., Lu, C.-T., and Zhang, P. 2001. Detecting graph-based spatial outliers : algorithms and applications (a summary of results). In Proceedings of the 7th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM Press, New York, NY, USA, 371-376.

[21] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark,R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. 2000. Phys-ioBank, PhysioToolkit, and PhysioNet : Components of a new research resource for com-plex physiologic signals. Circulation 101, 23, e215-e220.

[22] W.K. Lee, and S.J.Stolfo, 'A Data Mining Framework for Building Intrusion Detection Models,' In Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA : IEEE computer Society Press,1999, pp. 120-132.

[23] C. Elkan. (2007, Jan, 27). Results of the KDD'99 Knowledge Discovery Contest [Online]. Available : http ://www-cse.ucsd.edu/users/elkan clresults.html.

[24] X. Xu, X.N. Wang, ' Adaptive network intrusion detection method based on PCA and support vector machines,' Lecture Notes in Artificial Intelligence, ADMA 2005, LNAI 3584, 2005, pp. 696-703.

[25] Lerman Liran, ' Les systèmes de détection d'intrusion basés sur du machine learning 'Mémoire de Master, état de l'art. L'UNIVERSITE LIBRE DE BRUXELLES, Faculté des Sciences, Département d'Informatique. 2010.

[26] Calas G., ' Etudes des principaux algorithmes de data mining,' Article de recherche cole ingnieurs en informatique EPITF, France, 2006.

[27] G. A. Miller, ' The cognitive revolution : a historical perspective ' Trends in Cognitive Sciences, vol. 7, no. 3, pp. 141-144, Mar. 2003.

[28] Peddabachigari S, A. Abraham, C. Grosan, and J. Thomas, Modeling intrusion detection system using hybrid intelligent systems, Journal of Network and Computer Applications, vol. 30, no. 1, pp. 114-132, Jan. 2007.

[29] Gill Stéphane, ' Type d'attaques,' 2003.

[30] Gerphagnon Jean-Olivier et al, Attaques Informatique, Centro Brasileiro de Pesquisas Fisicas - CBPF/CNPq Coordena¸cao de Atividade Técnicas - CAT Rua Dr. Xavier Sigaud 150 22290-180 Rio de Janeiro - RJ - Brazil , CBPF-NT-007/00.

[31] Burgermeister David, Krier Jonathan, ' les systèmes de détection d'intrusions ' ,Developpez.com, Juillet 2006.

[32] J. Taylor, N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge UK, 2004.

[33] B. Lkopf, Kernel Principal Component Analysis Advances in Kernel Methods Support Vector Learning, Cambridge, UK, 1999.

[34] Raúl Morales, Felipe Valencia, Doris Sáez, Matías Lacalle, Supervisory Fuzzy Predictive Control for a Concentrated Solar Power Plant, Preprints of the 19th World Congress, Cape Town, South Africa, August 24-29,2014.

[35] J.Gertler,J.Romera,V.Puig,J.Quevedo, Détection des fuites et isolement dans les réseaux de distribution d'eau en utilisant l'analyse des composants principaux et les résidus structurés,CCFTS2010, 2010.

[36] catherin Zamora, juan manual Giraldo, Sylvain Leirens, Modèle de contrôle prédictif des réseaux de transport d'eau, 2010 IEEE.

[37] Demetrios G, Eliades and Marios M. Polycarpou, Optimisation multi-objectifs du placement des capteurs de qualité de l'eau dans les réseaux de distribution d'eau potable, Proceeding of the European Control Conference 2007, Kos, Greece, July 2-5, 2007.

[38] W.K. Lee, and S.J.Stolfo, 'A Data Mining Framework for Building Intrusion Detection Models,' In Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA : IEEE computer Society Press,1999, pp. 120-132.

[39] X. Xu, X.N. Wang, 'Adaptive network intrusion detection method based on PCA and support vector machines,' Lecture Notes in Artificial Intelligence, ADMA 2005, LNAI 3584, 2005, pp. 696-703