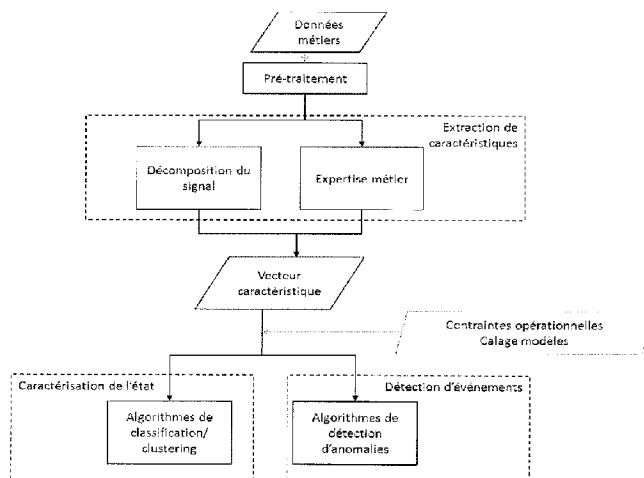




(86) Date de dépôt PCT/PCT Filing Date: 2015/07/23  
 (87) Date publication PCT/PCT Publication Date: 2016/01/28  
 (45) Date de délivrance/Issue Date: 2022/08/16  
 (85) Entrée phase nationale/National Entry: 2017/01/10  
 (86) N° demande PCT/PCT Application No.: IB 2015/055583  
 (87) N° publication PCT/PCT Publication No.: 2016/012972  
 (30) Priorité/Priority: 2014/07/25 (FR14 57209)

(51) Cl.Int./Int.Cl. *G01M 99/00* (2011.01)  
 (72) Inventeurs/Inventors:  
 CAMPAN, FRANCIS, FR;  
 DEMBELE, ABEL, FR;  
 CUSSONNEAU, GUILLAUME, FR  
 (73) Propriétaire/Owner:  
 SUEZ GROUPE, FR  
 (74) Agent: ANGLEHART ET AL.

(54) Titre : PROCÉDE POUR DETECTER DES ANOMALIES DANS UN RESEAU DE DISTRIBUTION, EN PARTICULIER D'EAU POTABLE  
 (54) Title: METHOD FOR DETECTING ANOMALIES IN A DISTRIBUTION NETWORK, IN PARTICULAR FOR DRINKING WATER



(57) **Abrégé/Abstract:**

Le réseau de distribution est équipé de capteurs. Les étapes sont: -acquérir pour chaque capteur une série temporelle de mesures physiques séparées par des intervalles de temps, -définir des fenêtres temporelles correspondant chacune à plusieurs intervalles de temps successifs, -extraire des caractéristiques opérationnelles de chaque série temporelle dans chaque fenêtre de temps, -former pour chaque fenêtre temporelle au moins un vecteur courant ayant pour coordonnées les caractéristiques opérationnelles, des données structurales relatives au réseau et des caractéristiques conjoncturelles propres à la fenêtre temporelle, -comparer le vecteur courant avec des vecteurs précédents, correspondant à des fenêtres temporelles précédentes, et dont les caractéristiques conjoncturelles et les données structurales sont semblables à celles du vecteur courant, et -signaler une anomalie dans des cas où le vecteur courant est significativement dissemblable desdits vecteurs précédents.

## (12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la  
Propriété Intellectuelle  
Bureau international(10) Numéro de publication internationale  
**WO 2016/012972 A1**(43) Date de la publication internationale  
28 janvier 2016 (28.01.2016)

WIPO | PCT

(51) Classification internationale des brevets :  
G01M 3/00 (2006.01)(21) Numéro de la demande internationale :  
PCT/IB2015/055583(22) Date de dépôt international :  
23 juillet 2015 (23.07.2015)

(25) Langue de dépôt : français

(26) Langue de publication : français

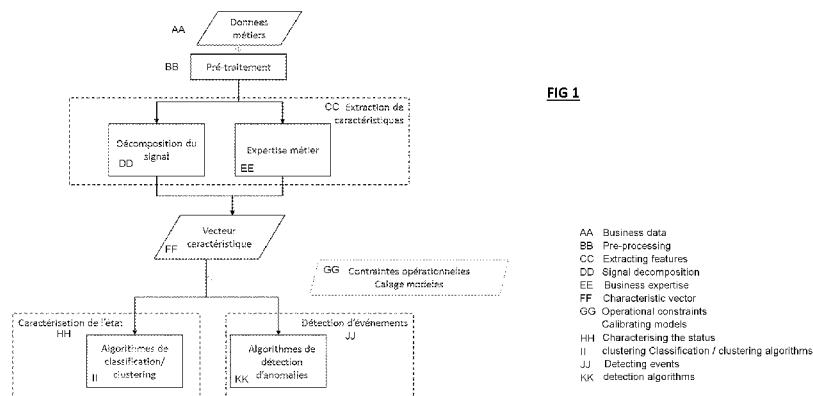
(30) Données relatives à la priorité :  
14 57209 25 juillet 2014 (25.07.2014) FR(71) Déposant : SUEZ ENVIRONNEMENT [FR/FR]; Tour  
CB21 - 16, Place de l'Iris, F-92040 Paris La Défense Cedex  
(FR).(72) Inventeurs : CAMPAN, Francis; 124, avenue François  
Molé, F-92160 Antony (FR). DEMBELE, Abel; 15, che-  
min des Plantes, F-69129 Villeurbanne (FR). CUSSON-  
NEAU, Guillaume; 73, rue La Fayette, F-75009 Paris  
(FR).(74) Mandataires : MOUGET-GONIOT, Claire et al.; 3, ave-  
nue Bugeaud, F-75116 Paris (FR).(81) États désignés (sauf indication contraire, pour tout titre  
de protection nationale disponible) : AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,  
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,  
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,  
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,  
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,  
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,  
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,  
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,  
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.(84) États désignés (sauf indication contraire, pour tout titre  
de protection régionale disponible) : ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,  
TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU,  
TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE,  
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,  
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,  
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, KM, ML, MR, NE, SN, TD, TG).

Publiée :

— avec rapport de recherche internationale (Art. 21(3))

(54) Title : METHOD FOR DETECTING ANOMALIES IN A DISTRIBUTION NETWORK, IN PARTICULAR FOR DRINKING WATER

(54) Titre : PROCÉDE POUR DÉTECTER DES ANOMALIES DANS UN RESEAU DE DISTRIBUTION, EN PARTICULIER D'EAU POTABLE



(57) Abstract : The invention relates to a distribution network provided with sensors. The method includes the following steps: acquiring for each sensor a time series of physical measurements separated by time intervals; defining time windows each corresponding to a plurality of consecutive time intervals; extracting operational features of each time series in each time window; forming for each time window at least one current vector having as coordinates the operational features, structural data relating to the network and temporary features specific to the time window; comparing the current vector with preceding vectors, corresponding to preceding time windows, and in which the temporal features and the structural data are similar to those of the current vector; and reporting an anomaly if the current vector differs significantly from said preceding vectors.

(57) Abrégé : Le réseau de distribution est équipé de capteurs. Les étapes sont: -acquérir pour chaque capteur une série temporelle de mesures

[Suite sur la page suivante]

**WO 2016/012972 A1**

---

physiques séparées par des intervalles de temps, -définir des fenêtres temporelles correspondant chacune à plusieurs intervalles de temps successifs, -extraire des caractéristiques opérationnelles de chaque série temporelle dans chaque fenêtre de temps, -former pour chaque fenêtre temporelle au moins un vecteur courant ayant pour coordonnées les caractéristiques opérationnelles, des données structurelles relatives au réseau et des caractéristiques conjoncturelles propres à la fenêtre temporelle, -comparer le vecteur courant avec des vecteurs précédents, correspondant à des fenêtres temporelles précédentes, et dont les caractéristiques conjoncturelles et les données structurelles sont semblables à celles du vecteur courant, et -signaler une anomalie dans des cas où le vecteur courant est significativement dissemblable desdits vecteurs précédents.

PROCEDE POUR DETECTER DES ANOMALIES DANS UN RESEAU DE  
DISTRIBUTION, EN PARTICULIER D'EAU POTABLE

5        Domaine de l'Invention

La présente invention concerne un procédé pour détecter des anomalies dans un réseau de distribution, en particulier distribution d'un fluide newtonien, encore plus particulièrement un réseau d'approvisionnement en eau potable.

10

La performance opérationnelle est au cœur de la gestion des systèmes d'approvisionnement en eau potable. Le niveau de performance peut être sensiblement amélioré grâce à des outils de détection et localisation d'anomalies hydrauliques sur les réseaux de transport et de distribution d'eau

15

potable. L'invention peut être utilisée pour tous les réseaux de fluides newtoniens, par exemple les réseaux urbains de chaleur et de climatisation.

20 On connaît pour la détection d'anomalies des méthodes de détection d'anomalies basées sur des techniques statistiques, telles que l'analyse du débit minimal d'un secteur du réseau, largement utilisée par les exploitants de réseaux d'eau potable. Ces méthodes ne permettent pas toujours de caractériser ni de localiser l'anomalie.

25

On connaît aussi des méthodes de détection basées sur des modèles hydrauliques. Ces méthodes basées sur des équations hydrauliques utilisent une modélisation du réseau, qui n'est pas toujours disponible. La détection et la localisation d'anomalie combinent des outils mathématiques et des mesures

30

physiques, ce qui implique la mise en place d'une densité d'instrumentation adéquate.

35

Le but de l'invention est de proposer un procédé de détection d'anomalies qui soit performant tout en étant peu exigeant en données d'entrée et relativement économique en termes de puissance de traitement informatique.

Suivant l'invention, le procédé pour détecter des anomalies dans un réseau de distribution, en particulier distribution d'un fluide newtonien, encore plus particulièrement un réseau d'approvisionnement en eau potable, le réseau de distribution étant équipé de capteurs, procédé dans lequel on acquiert pour  
5 chaque capteur une série temporelle de mesures physiques séparées par des intervalles de temps, est caractérisée par les étapes suivantes :

- définir des fenêtres temporelles correspondant chacune à plusieurs intervalles de temps,
- extraire des caractéristiques opérationnelles de chaque série temporelle  
10 dans chaque fenêtre de temps,
- former au moins un vecteur courant pour chaque fenêtre temporelle ayant pour coordonnées les caractéristiques opérationnelles, des données structurelles relatives au réseau et des caractéristiques conjoncturelles propres à la fenêtre temporelle,
- 15 - comparer le vecteur courant avec des vecteurs précédents, correspondant à des fenêtres temporelles précédentes, et dont les caractéristiques conjoncturelles et les données structurelles sont semblables à celles du vecteur courant,
- signaler une anomalie dans des cas où le vecteur courant est  
20 significativement dissemblable desdits vecteurs précédents.

Pour la comparaison du vecteur courant avec les vecteurs précédents, on définit un paramètre de sensibilité correspondant à un degré minimal de dissimilitude en dehors duquel le vecteur courant est qualifié de significativement dissemblable:  
25

L'intervalle de temps entre deux mesures d'une série temporelle est par exemple de quelques minutes, par exemple 3 mn.

Une fenêtre temporelle est typiquement d'une journée ou d'une semaine.  
30 Ainsi une série temporelle comprend typiquement de nombreuses mesures. On peut avantageusement mettre en œuvre le procédé simultanément pour des fenêtres temporelles de durées différentes.

Pendant une fenêtre temporelle, les paramètres mesurés par les capteurs,  
35 et constituant les séries temporelles, subissent des variations, par exemple en fonction de paramètres conjoncturels tels que l'heure, le jour de la

semaine, la saison, la météorologie, etc., de données structurelles du réseau, comme par exemple fermeture d'une vanne, arrêt d'une pompe, et aussi en fonction d'une ou plusieurs anomalies sur le réseau.

5 Suivant l'invention on forme un vecteur ayant des dimensions représentatives des séries temporelles, des dimensions représentatives des paramètres conjoncturels, et des dimensions représentatives des données structurelles.

10 L'idée qui est à la base de l'invention est que si les dimensions représentatives des paramètres conjoncturels et des données structurelles ont été pertinemment choisies, les vecteurs ayant sensiblement les mêmes coordonnées pour les dimensions conjoncturelles et structurelles devraient aussi avoir sensiblement les mêmes coordonnées pour les dimensions  
15 opérationnelles. Dans le cas contraire on signale une anomalie.

Compte-tenu du nombre typiquement grand des valeurs numériques dans chaque série temporelle, le procédé devrait traiter des vecteurs ayant un très grand nombre de dimensions si toutes ces valeurs numériques  
20 devenaient chacune une coordonnée d'une dimension respective du vecteur. Ainsi, une autre idée qui est à la base de l'invention consiste à extraire des caractéristiques de chaque série temporelle, et ensuite à utiliser ces caractéristiques pour des coordonnées du vecteur.

25 Les caractéristiques extraites des séries temporelles peuvent comprendre des maximums, des minimums et/ou des moyennes, et/ou encore des fréquences élémentaires, obtenues en particulier par décomposition des séries temporelles en séries de Fourier. Certains paramètres conjoncturels ou données structurelles, par exemple la température ambiante ou l'état  
30 d'une vanne, peuvent également être disponibles sous forme de série temporelle que l'on traite pour en extraire des caractéristiques qui constitueront des coordonnées du vecteur en lieu et place des mesures ou données brutes.

35 Pour l'extraction de caractéristiques, on peut même agréger des séries temporelles de mesures. On peut par exemple faire l'addition des

consommations connues d'après la télé-relève, et obtenir ainsi une mesure unique totale, ou encore une série temporelle des consommations totales, par exemple une mesure par jour dans le cas d'une fenêtre temporelle d'une semaine.

5

De manière avantageuse, certaines caractéristiques conjoncturelles sont au moins en partie basées sur des hypothèses issues de l'expérience. Par exemple, on peut disposer des tendances sur la consommation d'eau chez un particulier en fonction de l'heure, du jour de la semaine, de la saison, de la température extérieure, de la pluviométrie, du nombre d'occupants de l'habitation etc.

10

Dans un mode de mise en œuvre avantageux, en cas d'anomalie, on compare le vecteur courant avec au moins un vecteur précédent ayant des caractéristiques conjoncturelles et des données structurelles semblables, et des caractéristiques opérationnelles aussi proches que possible du vecteur courant, et on signale au moins une caractéristique opérationnelle pour laquelle le vecteur courant présente un grand écart avec ledit au moins un vecteur précédent.

15

20

De préférence, on traite par un logiciel de caractérisation d'anomalies les situations ayant donné lieu à la signalisation d'au moins une anomalie. Un tel logiciel fonctionne de façon bien plus efficace lorsqu'il ne traite que des situations triées comme anormales avec en outre des informations déjà disponibles sur les paramètres (ou coordonnées de vecteur) présentant une valeur anormale.

25

Le résultat, anomalie d'un vecteur ou anomalie de réseau déterminée après traitement par un logiciel de caractérisation comme indiqué ci-dessus, est de préférence fourni par référence à une échelle de gravité de l'anomalie.

30

Ayant quantifié la gravité d'anomalie, en cas de plusieurs anomalies simultanées, on peut prioriser les anomalies les unes par rapport aux autres en fonction de leur urgence et/ou en fonction de l'ampleur de l'intervention corrective correspondante.

35

Dans un mode de mise en œuvre préféré, un vecteur exempt d'anomalies est classé dans un même compartiment de mémoire que d'autres vecteurs sensiblement équipollents, et chaque fois qu'un vecteur courant a été calculé on recherche le compartiment de mémoire contenant les vecteurs antérieurs qui lui ressemblent le plus, et on ne compare le vecteur courant qu'avec les vecteurs précédents de ce compartiment. On réduit ainsi la puissance de traitement nécessaire et le délai d'obtention des résultats.

On crée un nouveau compartiment de mémoire pour un vecteur dont les coordonnées conjoncturelles et/ou les coordonnées correspondant à des données structurelles du réseau ne correspondent à aucun compartiment existant. Par exemple une météorologie exceptionnellement fraîche pour un mois d'été peut donner lieu à la création d'un nouveau compartiment.

Il est avantageux d'analyser l'évolution dans le temps des vecteurs d'un même compartiment, et de fournir des informations sur l'évolution du réseau. On peut ainsi, par exemple, révéler des fuites initialement faibles mais ayant tendance à s'aggraver et pouvant devenir importantes, alors que la simple recherche du vecteur le plus semblable dans le compartiment ne révélera pas la fuite croissante, puisque le vecteur précédent le plus récent paraîtra très semblable et conduira à conclure à l'absence d'anomalie.

Dans une version encore plus perfectionnée, on compare les évolutions dans les différents compartiments, et on fournit des informations qui différencient l'évolution de l'état du réseau et l'évolution des conséquences des paramètres conjoncturels sur le réseau. Par exemple une consommation qui n'augmente que les jours de forte chaleur indique une évolution des habitudes des consommateurs plutôt qu'un réseau de plus en plus grevé de fuites.

Pour alléger le traitement et affiner les résultats, il est très avantageux de fournir pour chaque fenêtre temporelle plusieurs vecteurs correspondant chacun à un sous-réseau respectif faisant partie du réseau. Ainsi on traite des vecteurs plus petits et on se donne des chances supplémentaires de localiser plus facilement une anomalie.

5 En cas de détection d'une anomalie on analyse comparativement le vecteur courant avec des vecteurs récents comparablement anormaux, pour fournir des indications sur une vitesse d'évolution de l'anomalie et/ou un lien de l'anomalie avec au moins un paramètre conjoncturel et/ou une donnée structurelle.

10 En cas d'anomalie, on peut également rechercher dans les vecteurs précédents des anomalies comparables ayant donné lieu à un diagnostic, pour fournir un pré-diagnostic de la cause de l'anomalie courante.

Lors d'une étape d'initialisation, on peut avantageusement selon l'invention charger une mémoire avec des vecteurs reconstitués d'après des archives relatives au réseau.

15 Les composantes d'un vecteur incluent de préférence au moins une composante relative aux plaintes des consommateurs desservis par le réseau, par exemple sur le débit, la pression, le goût etc..

20 D'autres particularités et avantages de l'invention ressortiront encore de la description ci-après, relative à des exemples non-limitatifs, et en référence à la figure 1 qui représente un organigramme des étapes principales dans l'exemple décrit.

### **Observation préliminaire**

25 La description qui suit vaut description de toute particularité qu'elle contient, qu'elle soit prise isolément des autres particularités même faisant partie du même paragraphe ou de la même phrase, et vaut description de toute combinaison de telles particularités, dès lors qu'une telle particularité ou  
30 combinaison de particularités est distinctive de l'état de la technique et offre un effet technique, que ce soit seule ou en combinaison avec des concepts présentés ci-dessus, et que cette particularité soit exprimée dans les termes même de cette description ou dans des termes plus ou moins généralisés.

## Définitions

- Entité : Réseau d'eau potable ou composante du système d'approvisionnement en eau potable, par exemple les secteurs hydrauliques ou les appareils de mesure/capteurs. A  
5 une entité sont associées une ou plusieurs séries temporelles.
- Série temporelle : séquence finie de données scalaires indexées par le temps, généralement espacées d'une durée constante.
- Classification : Sans autre précision, désigne le processus aboutissant à assigner à un état d'une entité une classe connue donnant lieu à des actions  
10 prédéterminées de la part de l'opérateur de réseau.
- Génération de classification (« Clustering ») : Sans autre précision, désigne le processus aboutissant à associer à un état d'une entité un groupe d'états antérieurs afin d'en déterminer le caractère anormal ou nouveau.
- 15 Mesure : C'est l'estimation de la valeur d'une grandeur en unité correspondant, ces valeurs formant une série temporelle fournie par un système d'acquisition de données. Une mesure est associée à un composant élémentaire particulier du réseau d'eau potable (arc ou tronçon pour une mesure de  
20 débit, nœud pour une mesure de pression, réservoir pour une mesure de niveau).
- Données de télé-relève : série d'index de consommation pour un compteur, mesurés à une périodicité donnée, et télé transmise par exemple au moins une fois par jour.
- 25 Caractéristique : (« feature ») scalaire ou vecteur constituant une information significative pour le processus étudié, ici l'état du réseau sur une fenêtre temporelle donnée.

## Principes

30

Dans ce mode réalisation l'invention met en œuvre les principes suivants :

- Le prétraitement du signal par les méthodes de l'état de l'art permettant de le compléter et de le nettoyer du bruit.
- La production d'un vecteur de caractéristiques de l'état de  
35 fonctionnement d'une entité pour une fenêtre temporelle donnée. Celle-ci se fait par agrégation :

- de caractéristiques extraites principalement des séries temporelles de l'entité par des méthodes de décomposition/traitement du signal,
  - de caractéristiques issues d'indicateurs de performance métier et de données conjoncturelles.
- 5
- L'utilisation d'algorithmes de classification et de « clustering », issus des techniques d'auto-apprentissage machine (« machine learning »), appliqués aux vecteurs caractéristiques des entités pour une fenêtre temporelle donnée. Ceci permet ensuite la transcription des états des entités (tels que représentés par les vecteurs caractéristiques) en des catégories connues des exploitants de réseaux, afin de classer les situations et de prioriser les actions correctives à mener. Les algorithmes de classification/clustering sont calés sur des fonctions-objectifs et des contraintes opérationnelles de gestion du réseau, par exemple l'effectif disponible pour une opération d'entretien simple, le délai pour disposer d'une équipe pour une opération plus complexe, le délai entre la date d'apparition d'une anomalie et la date de sa détection, etc.
- 10
- L'utilisation optionnelle de données de contexte influençant le fonctionnement de l'entité, pour préciser la description de son état.
- 15
- L'utilisation d'algorithmes de détection/évaluation (« detection/scoring ») d'anomalies issus des techniques d'autoapprentissage machine (« machine learning »), calées/optimisées à l'aide de critères répondant aux contraintes opérationnelles des exploitants de réseau. Ces algorithmes sont appliqués aux vecteurs de caractéristiques.
- 20

25

Données d'entrée et paramétrage : trois ensembles de données :

Les données structurelles sont constituées des données descriptives de l'infrastructure du réseau et des équipements installés (vannes, capteurs, pompes, etc.).

30 Les paramètres de mise en œuvre de la méthode sont calés de manière automatique dans une phase préparatoire, elle-même automatisée afin de permettre un recalage lorsque le système détecte une perte de performance ou une évolution de son infrastructure.

Les données opérationnelles sont issues de l'ensemble des systèmes de mesure présents sur le réseau. Lorsqu'elles sont disponibles ces données incluront, entre autres, aussi les données de consommation télé-

35

relevées, les plaintes clientèles et les interventions ayant une incidence sur le comportement du réseau.

### **Méthodes élémentaires** (Voir Figure 1)

5

#### Prétraitement des séries temporelles des mesures

Ces méthodes permettent d'obtenir des séries temporelles des mesures éventuellement complétées, lissées/nettoyées du bruit. Les séries temporelles résultantes sont alors prêtes à être utilisées comme données d'entrée des algorithmes d'extraction de caractéristiques.

Des combinaisons des séries en signaux agrégés sont également réalisées. Par exemple la somme algébrique des séries temporelles de débits d'entrée/sortie d'un secteur hydraulique est convertie en série de consommations du secteur hydraulique.

Les séries peuvent également être transformées (par exemple centrées réduites) pour les besoins de certains des algorithmes utilisés dans les phases d'extraction de caractéristiques.

#### 20 Extraction des caractéristiques issues de la décomposition du signal

Cette méthode consiste à utiliser les séries temporelles précédemment nettoyées/lissées comme décrit ci-dessus, pour en extraire l'information pertinente permettant de caractériser de manière opérationnelle l'état d'une entité. Il s'agit de produire les informations résumant la structure des séries (décomposition de Fourier, en ondelettes, en composantes principales, ...) tout en en diminuant la dimension afin de concentrer la partie pertinente des signaux.

Les différentes bases de décomposition sont évaluées de manière périodique afin de suivre les performances des algorithmes et de mettre à jour les bases quand celles-ci sont jugées trop peu parcimonieuses (supervision).

Les sorties de ces algorithmes sont donc des vecteurs de caractéristiques résumant chaque signal dans une fenêtre temporelle définie par l'exploitant de réseau.

35

## Extraction des caractéristiques métier

La construction des caractéristiques métier d'un site est automatisée et s'appuie sur :

- 5
- un référentiel de caractéristiques établies à partir de l'expérience de réseaux de différents types,
  - une observation de la structure et du comportement du réseau ciblé.

10 La construction peut s'appuyer par exemple sur un calcul des minimums d'un paramètre sur une fenêtre temporelle donnée, l'observation de la périodicité des cycles de remplissage/vidange des réservoirs, le niveau de consommation moyen des compteurs télé-relevés en fonction de leur segment de consommateur, etc..

15 Algorithmes d'évaluation de l'état de fonctionnement de l'entité – régime de fonctionnement

20 Selon une particularité importante de l'invention, pour décrire l'état de fonctionnement du réseau, on combine des données issues de l'expertise métier avec celles issues des méthodes de décomposition du signal. Celles-ci viennent s'enrichir mutuellement en donnant en sortie pour chaque entité un vecteur caractérisant complètement un état de fonctionnement de cette entité pour le réseau considéré.

25 A l'aide des vecteurs de caractéristiques, l'état de fonctionnement de l'entité considérée à un temps donné dans une fenêtre temporelle donnée, ou « régime », peut être caractérisé. Il peut alors être comparé aux états antérieurs, ou aux états des autres entités, et classé selon des critères opérationnels impliquant des actions à engager. Pour réaliser ce classement, les  
30 outils de machine learning tels que la classification et le clustering peuvent être mis en œuvre.

Dans le cas de la classification, un processus d'apprentissage aura préalablement permis, à l'aide de données historiques marquées avec les  
35 différents états possibles et constituant des groupes, d'entraîner une fonction

discriminante (« classifieur »), qui permet de donner automatiquement le groupe auquel le nouvel état appartient.

5 Cette fonction discriminante peut être obtenue par optimisation d'un critère de qualité, par exemple une fonction d'inertie en mode non supervisé ou le proxy d'une erreur de classification en mode supervisé.

10 Dans le cas d'une absence de données historiques marquées, les techniques de génération de classification (« clustering ») permettent de grouper les états selon des critères de ressemblance et ainsi de discriminer ceux qui appartiennent aux catégories les moins représentées.

15 Celles-ci présentent un intérêt pour l'opérateur de réseau puisqu'elles indiquent un comportement qui sort de l'ordinaire. L'opérateur peut alors focaliser son attention sur cette entité.

#### Algorithmes de détection d'anomalies

20 Aux vecteurs de caractérisation d'état sont appliqués des algorithmes de détection d'anomalies. Il s'agit de caractériser un événement en cours ou récemment terminé. Le type d'événement correspond à une catégorie d'événements suivis par les opérateurs de réseaux d'eau potable (fuites, chute de pression, panne de capteur, anomalie de consommation, etc.).

25 Les algorithmes de détection sont lancés en parallèle et leurs résultats sont agrégés pour la discrimination des entités présentant un comportement anormal. Ils ont été préalablement calés sur des jeux de données historiques, afin d'adapter leurs paramétrages aux contraintes opérationnelles du réseau.

#### 30 **Particularités corollaires**

L'ensemble constitué par les composants exposés plus haut peut être connecté au système d'information technique de l'opérateur d'un système d'approvisionnement en eau potable. Chaque source de données disponible est  
35 alors connectée à l'algorithme d'extraction de caractéristiques qui lui est dédié. L'ensemble peut être activé régulièrement, selon la fréquence d'acquisition des

données. La fenêtre temporelle sur laquelle se fait l'analyse est réglable par l'utilisateur. Il est néanmoins pertinent d'utiliser en première approche des fenêtres temporelles de 24 heures et 7 jours.

- 5 Dans ces conditions la caractérisation de l'état de fonctionnement du réseau et la détection d'anomalies associée sont beaucoup plus pertinentes qu'avec les méthodes habituellement utilisées, et l'efficacité opérationnelle en est considérablement améliorée.
- 10 Dans la phase de diagnostic des événements passés d'un réseau, à des fins de bilan d'exploitation par exemple, la méthode permet de gagner un temps considérable en discriminant les situations passées événementielles.

15 La combinaison des caractéristiques liées aux signaux et des caractéristiques métier améliore la robustesse de la détection d'anomalie.

Le réglage de la sensibilité de l'algorithme, c'est-à-dire l'équilibre entre le nombre de détections justes et le nombre d'anomalies pour chaque entité, permet, par exemple, d'adapter la détection de chaque type d'anomalie à la  
20 capacité de l'exploitant à planifier et engager les actions correctives.

### **Exemple de réalisation : Exemple chlore et décomposition de signal**

L'utilisation d'une méthode de décomposition du signal, telle que décomposition  
25 en ondelettes ou de Fourier, sur des signaux relatifs à un réseau d'eau potable, permet d'isoler les différentes composantes (intra et inter jour) de ces signaux. Un algorithme d'analyse cognitive de ces composantes, en relation avec l'expertise métier, aboutit à identifier le domaine de définition du fonctionnement normal du réseau. Ceci permet, lorsque de nouveaux signaux  
30 sont disponibles, de détecter les changements significatifs dans la nature de ces composantes.

Ainsi, un écart significatif observé peut être interprété comme indicateur de comportement anormal. Un second ensemble d'algorithmes de recherche peut  
35 alors être exécuté. Par exemple, un algorithme de classification sur la base des

WO 2016/012972

PCT/IB2015/055583

vecteurs de caractéristiques (incluant les plaintes clients par exemple) peut aider à définir le niveau de risque atteint.

5 Cette approche peut être appliquée pour la surveillance de la qualité de l'eau en utilisant l'ensemble des points de mesure disponibles sur un réseau d'eau potable afin d'identifier les structures cachées et détecter, par exemple, des anomalies sur la concentration résiduelle de chlore en conjonction ou pas avec d'autres paramètres qualité. Les plaintes clients utilisées sont alors relatives au goût de l'eau.

10

## RENDICATIONS

- 5 1. Procédé pour détecter des anomalies dans un réseau de distribution, en  
particulier distribution d'un fluide newtonien, encore plus particulièrement  
un réseau d'approvisionnement en eau potable, le réseau de distribution  
étant équipé de capteurs, procédé dans lequel on acquiert pour chaque  
10 capteur une série temporelle de mesures physiques séparées par des  
intervalles de temps, caractérisé par les étapes suivantes :
- définir des fenêtres temporelles correspondant chacune à plusieurs  
intervalles de temps,
  - extraire des caractéristiques opérationnelles de chaque série temporelle  
dans chaque fenêtre de temps,
  - 15 - former pour chaque fenêtre temporelle au moins un vecteur courant  
ayant pour coordonnées les caractéristiques opérationnelles, des données  
structurelles relatives au réseau et des caractéristiques conjoncturelles  
propres à la fenêtre temporelle,
  - comparer le vecteur courant avec des vecteurs précédents,  
20 correspondant à des fenêtres temporelles précédentes, et dont les  
caractéristiques conjoncturelles et les données structurelles sont  
semblables à celles du vecteur courant, et
  - signaler une anomalie dans des cas où le vecteur courant est  
significativement dissemblable desdits vecteurs précédents.
- 25 2. Procédé selon la revendication 1, dans lequel pour la comparaison du  
vecteur courant avec les vecteurs précédents on règle une sensibilité  
définissant un degré minimal de dissimilitude pour que le vecteur courant  
soit qualifié de significativement dissemblable.
- 30 3. Procédé selon la revendication 1 ou 2, dans lequel pour l'extraction de  
caractéristiques, on agrège des séries temporelles de mesures.
4. Procédé selon l'une des revendications 1 à 3, dans lequel les  
caractéristiques conjoncturelles sont au moins en partie basées sur des  
hypothèses issues de l'expérience.
- 35 5. Procédé selon l'une des revendications 1 à 4, dans lequel en cas  
d'anomalie, on compare le vecteur courant avec au moins un vecteur

- précédent ayant des caractéristiques conjoncturelles semblables et des caractéristiques opérationnelles aussi proches que possible du vecteur courant, et on signale au moins une caractéristique opérationnelle pour laquelle le vecteur courant présente un grand écart avec ledit au moins un vecteur précédent.
- 5
6. Procédé selon l'une des revendications 1 à 5, dans lequel on fournit le résultat par référence à une échelle de gravité de l'anomalie.
7. Procédé selon l'une des revendications 1 à 6, dans lequel en cas de plusieurs anomalies simultanées on priorise les anomalies en fonction de leur urgence et/ou en fonction de l'ampleur de l'intervention corrective correspondante.
- 10
8. Procédé selon l'une des revendications 1 à 7, dans lequel un vecteur exempt d'anomalies est classé dans un même compartiment de mémoire que d'autres vecteurs sensiblement équipollents, et chaque fois qu'un vecteur courant a été calculé on recherche le compartiment de mémoire contenant les vecteurs antérieurs qui lui ressemblent le plus, et on ne compare le vecteur courant qu'avec les vecteurs précédents de ce compartiment.
- 15
9. Procédé selon la revendication 8, dans lequel on crée un nouveau compartiment de mémoire pour un vecteur dont les coordonnées conjoncturelles et/ou les coordonnées correspondant à des données structurelles du réseau ne correspondent à aucun compartiment existant.
- 20
10. Procédé selon l'une des revendications 1 à 9, dans lequel on analyse l'évolution dans le temps des vecteurs d'un même compartiment, et on fournit des informations sur l'évolution du réseau.
- 25
11. Procédé selon la revendication 10, dans lequel on compare les évolutions dans les différents compartiments, et on fournit des informations qui différencient l'évolution de l'état du réseau et l'évolution des conséquences des paramètres conjoncturels sur le réseau.
- 30
12. Procédé selon l'une des revendications 1 à 11, dans lequel on fournit pour chaque fenêtre temporelle plusieurs vecteurs correspondant chacun à un sous-réseau respectif faisant partie du réseau.
13. Procédé selon l'une des revendications 1 à 12, dans lequel les caractéristiques extraites des séries temporelles comprennent des maximums, des minimums et/ou des moyennes.
- 35

14. Procédé selon l'une des revendications 1 à 13, dans lequel les caractéristiques extraites des séries temporelles comprennent des fréquences élémentaires, obtenues en particulier par décomposition des séries temporelles en séries de Fourier.
- 5 15. Procédé selon l'une des revendications 1 à 14, dans lequel en cas de détections d'une anomalie on analyse comparativement le vecteur courant avec les vecteurs précédents pour fournir des indications sur une vitesse d'évolution de l'anomalie et/ou un lien de l'anomalie avec au moins un paramètre conjoncturel et/ou une donnée structurelle.
- 10 16. Procédé selon l'une des revendications 1 à 15, dans lequel lors d'une étape d'initialisation on charge une mémoire avec des vecteurs reconstitués d'après des archives relatives au réseau.
- 15 17. Procédé selon l'une des revendications 1 à 16, dans lequel on le met en œuvre simultanément pour des fenêtres temporelles de durées différentes.
18. Procédé selon l'une des revendications 1 à 17, dans lequel les composantes d'un vecteur incluent les plaintes des consommateurs desservis par le réseau.
- 20 19. Procédé selon l'une des revendications 1 à 18, dans lequel on traite par un logiciel de caractérisation d'anomalies les situations ayant donné lieu à la signalisation d'au moins une anomalie.
- 25 20. Procédé selon l'une des revendications 1 à 19, dans lequel en cas d'anomalie, on recherche dans des vecteurs précédents des anomalies comparables ayant donné lieu à un diagnostic, pour fournir pour la cause de l'anomalie courante un pré-diagnostic semblable au diagnostic antérieur.

1/1

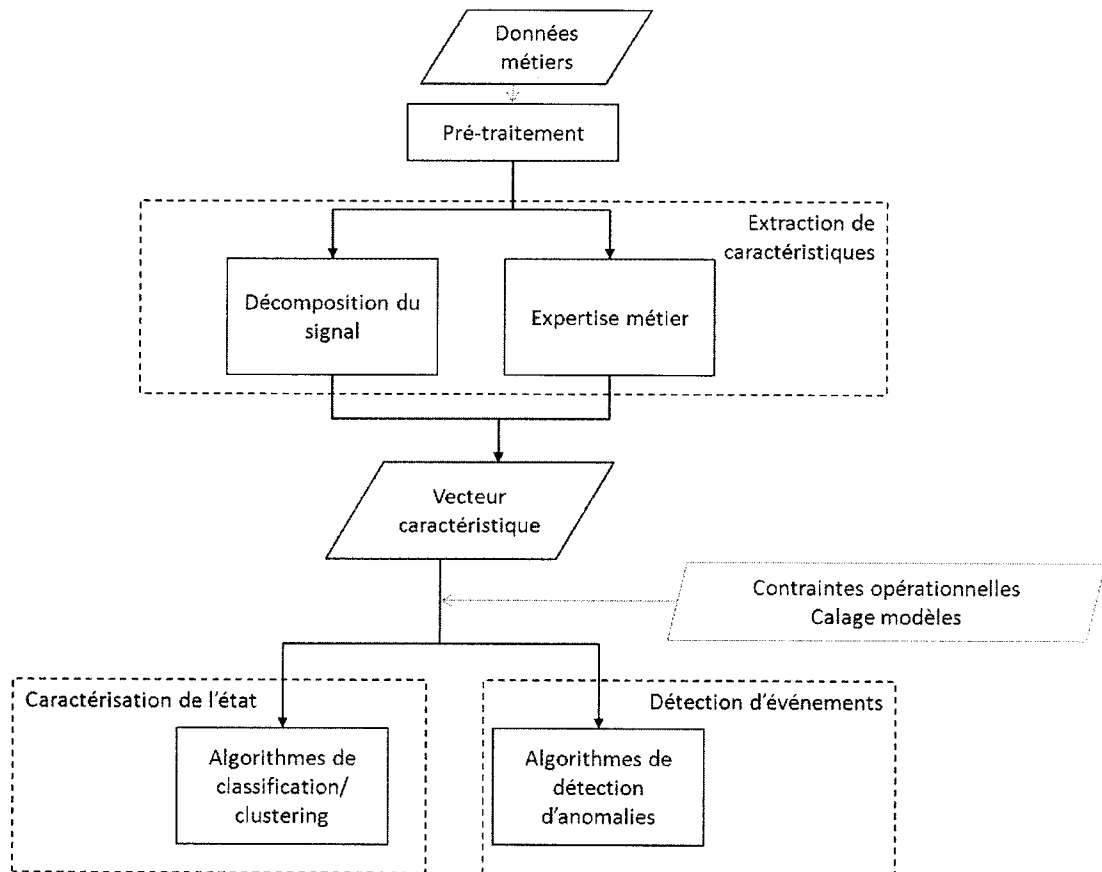


FIG 1

