

Plan de thèse

December 20^e 2017

Contents

1	Introduction Générale	5
2	Contexte de l'étude	7
2.1	État de l'art: reconstruction du graphe par les incidents	7
2.2	Problématique	7
3	Réseau de flots et Mesures	9
3.1	Réseau électrique d'un datacenter: un graphe de flots	9
3.2	Modélisation du réseau électrique	9
3.2.1	Grandeurs physiques	9
3.2.2	Mesures physiques : des Series Temporelles	9
3.2.3	Contraintes dans un reseau électrique : Loi de Kirchhoff	9
4	Corrélation de mesures	11
4.1	Analyse de séries temporelles	11
4.1.1	Séries Temporelles : domaines d'analyses	12
4.1.2	Pourquoi analyser nos mesures comme des series temporelles	12
4.2	Étude de l'art sur la corrélation de séries temporelles	12
4.2.1	Corrélation sur les séries entières: {W,D}DTW, TWE, MSM, CID, DTDc, COTE	12
4.2.1.1	Dynamic Time Warping: weight WDTW, derivative DDTW	12
4.2.1.2	Time Warp Edit TWE	12
4.2.1.3	Move-split-merge MSM	12
4.2.1.4	Complexity invariant Distance CID	12
4.2.1.5	Elastic Ensemble EE	12
4.2.1.6	Collective of Transform Ensembles COTE	12
4.2.1.7	12
4.2.2	Corrélation par intervalles: TSF, TSBF, LPS	12
4.2.2.1	Time Series Forest TSF	12
4.2.2.2	Time Series Bag of Features TSBF	12
4.2.2.3	Learned Pattern Similarity LPS	12
4.2.3	Corrélation par parties significatives(shapelets): FS, ST, LS	12
4.2.3.1	Fast Shapelets FS	12
4.2.3.2	Shapelet Transform ST	12
4.2.3.3	Learned Shapelets LS	12
4.2.4	Corrélation par agrégation des features: BOP, SAXVSM, BOSS, DTW _F	12
4.2.4.1	Bag of Pattern BOP	12

4.2.4.2	Symbolic Aggregate approXimation Vectr Space Model SAXVSM . . .	12
4.2.4.3	Bag of Symbolic Fourier Approximation (SFA) symbols BOSS . . .	12
4.2.4.4	Dynamic Time Warping Features DTW_F	12
4.3	Proposition d'une méthode de corrélation de mesures électriques	12
4.4	Formalisation et calcul de la Matrice de corrélation	12
4.5	Cas des mesures du datacenter Champlan	12
4.5.1	Description du datacenter Champlan	12
4.5.2	Matrice de corrélation : corrélation fausses négatives et fausses positives . .	12
5	Découverte de topologie par les corrélations	13
5.1	Description du problème	13
5.1.1	Propriétés de la matrice de corrélation	13
5.1.2	Définition du problème	13
5.1.3	Définition de la line-couverture	13
5.2	Algorithmes de découverte de topologie	13
5.2.1	Algorithme de couverture	13
5.2.2	Algorithme de correction	13
5.2.3	Complexité des algorithmes	13
5.3	Détermination de la topologie du réseau énergétique	13
5.4	Correction particulière de Graphes : Graphes lourdes	13
6	Evaluation des algorithmes sur des données theoriques et réelles	15
6.1	Simulation des algorithmes sur des réseaux théoriques	15
6.1.1	Objectifs et définitions	15
6.1.2	Données: Génération aléatoires de graphes	15
6.1.3	Génération de line graphs sous jacent aux réseaux de flots non orienté . . .	15
6.1.4	Prise en compte de l'erreur de corrélation dans la matrice $matE$	15
6.1.5	Résultats	15
6.1.5.1	Distribution de la méthode de permutation aléatoire	15
6.1.5.2	Relation entre la distance-line et la distance de Hamming	15
6.1.5.3	Comparaison des méthodes de correction	15
6.1.5.4	Influence des erreurs de corrélations sur les distributions	15
6.1.5.5	Impact de la fonction de coût sur les distributions	15
7	Conclusions	17
8	Perpectives	19

Chapter 1

Introduction Générale

Chapter 2

Contexte de l'étude

2.1 État de l'art: reconstruction du graphe par les incidents

Dans cette partie, nous utilisons le cas du réseau électrique de ERDF, découvert à partir des incidents. Nous décrivons les étapes pour parvenir à ce resultat.

2.2 Problématique

Nous définirons le problème proxi-line qui est la distance de Hamming minimale selon certaines conditions.

Chapter 3

Réseau de flots et Mesures

3.1 Réseau électrique d'un datacenter: un graphe de flots

Dans cette section, nous allons decire un reseau de datacenter et faire le lien entre ce reseau et un graphe de flots.

3.2 Modélisation du réseau électrique

Nous allons decire les grandeurs presentes dans le datacenter et leurs mesures associées. Les mesures des grandeurs physiques constituent le flot dans le graphe. Nous parlerons des contraintes liées aux flots dans le reseau électriques. Ces contraintes ont le nom de loi de conservation ou de kirrchhoff(loi des noeuds(I,P) et des mailles (U)).

3.2.1 Grandeurs physiques

3.2.2 Mesures physiques : des Series Temporelles

nous definirons une serie temporelle, donnerons ses proprietes et ses avantages.

3.2.3 Contraintes dans un reseau électrique : Loi de Kirchhoff

Chapter 4

Corrélation de mesures

4.1 Analyse de séries temporelles

Ici nous parlerons des types d'analyses qu'on réalise avec les series temporelles. Nous indiquerons le type d'analyses que nous souhaiterons faire sur nos mesures et le domaine qui s'y rattache.

4.1.1 Séries Temporelles : domaines d'analyses

4.1.2 Pourquoi analyser nos mesures comme des series temporelles

4.2 Étude de l'art sur la corrélation de séries temporelles

4.2.1 Corrélation sur les séries entières: {W,D}DTW, TWE, MSM, CID, DTDc, COTE

4.2.1.1 Dynamic Time Warping: weight WDTW, derivative DDTW

4.2.1.2 Time Warp Edit TWE

4.2.1.3 Move-split-merge MSM

4.2.1.4 Complexity invariant Distance CID

4.2.1.5 Elastic Ensemble EE

4.2.1.6 Collective of Transform Ensembles COTE

4.2.1.7

4.2.2 Corrélation par intervalles: TSF, TSBF, LPS

4.2.2.1 Time Series Forest TSF

4.2.2.2 Time Series Bag of Features TSBF

4.2.2.3 Learned Pattern Similarity LPS

4.2.3 Corrélation par parties significatives(shapelets): FS, ST, LS

4.2.3.1 Fast Shapelets FS

4.2.3.2 Shapelet Transform ST

4.2.3.3 Learned Shapelets LS

4.2.4 Corrélation par agrégation des features: BOP, SAXVSM, BOSS, DTW_F

4.2.4.1 Bag of Pattern BOP

4.2.4.2 Symbolic Aggregate approXimation Vectr Space Model SAXVSM

4.2.4.3 Bag of Symbolic Fourier Approximation (SFA) symbols BOSS

4.2.4.4 Dynamic Time Warping Features DTW_F

4.3 Proposition d'une méthode de corrélation de mesures électriques

4.4 Formalisation et calcul de la Matrice de corrélation

4.5 Cas des mesures du datacenter Champlan

4.5.1 Description du datacenter Champlan

Chapter 5

Découverte de topologie par les corrélations

5.1 Description du problème

5.1.1 Propriétés de la matrice de corrélation

5.1.2 Définition du problème

5.1.3 Définition de la line-couverture

5.2 Algorithmes de découverte de topologie

5.2.1 Algorithme de couverture

5.2.2 Algorithme de correction

5.2.3 Complexité des algorithmes

5.3 Détermination de la topologie du réseau énergétique

5.4 Correction particulière de Graphes : Graphes Iourtes

Chapter 6

Evaluation des algorithmes sur des données theoriques et réelles

6.1 Simulation des algorithmes sur des réseaux théoriques

6.1.1 Objectifs et définitions

6.1.2 Données: Génération aléatoires de graphes

6.1.3 Génération de line graphes sous jacent aux réseaux de flots non orienté

6.1.4 Prise en compte de l'erreur de corrélation dans la matrice $matE$

6.1.5 Résultats

6.1.5.1 Distribution de la méthode de permutation aléatoire

6.1.5.2 Relation entre la distance-line et la distance de Hamming

6.1.5.3 Comparaison des méthodes de correction

6.1.5.4 Influence des erreurs de corrélations sur les distributions

6.1.5.5 Impact de la fonction de coût sur les distributions

Chapter 7

Conclusions

Chapter 8

Perspectives

Contents