

Apprentissage de motifs temporels, multidimensionnels et hétérogènes

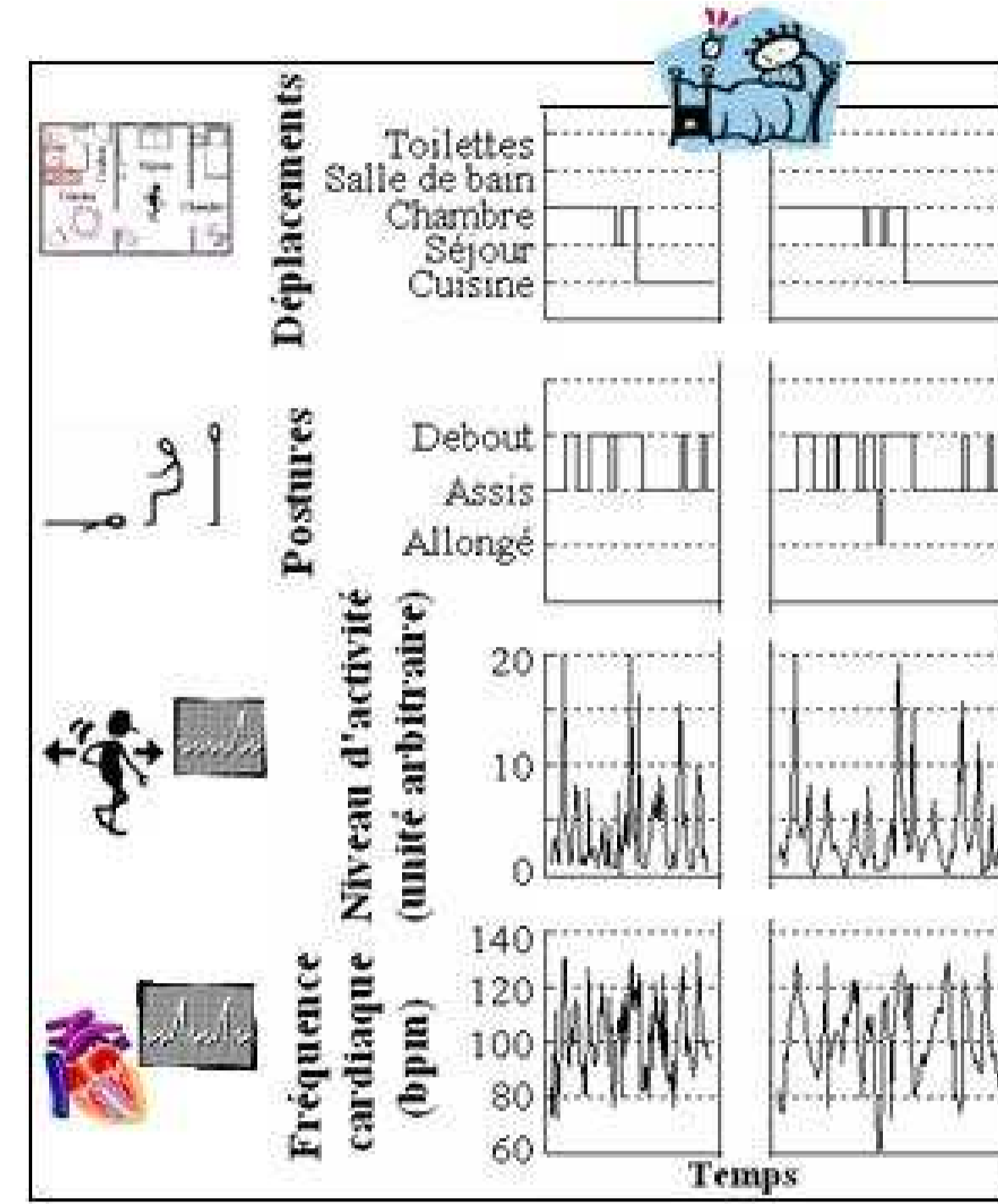
Application à la télésurveillance de personnes à domicile

Florence Duchêne¹, Catherine Garbay¹ et Vincent Rialle^{1,2}

¹ Laboratoire TIMC-IMAG, Faculté de médecine de Grenoble, France - ² SIIM, CHU de Grenoble, France
 Florence.Duhene@sophia.inria.fr, Catherine.Garbay@imag.fr, Vincent.Rialle@imag.fr

1. Introduction

- Contexte** → **Système d'aide à la décision pour la télésurveillance**
 - Observation dans le temps par un ou plusieurs capteurs
 - Objectif : Détection de situations indésirables (ex.: baisse d'activité d'une personne fragile ou isolée)
- Objectif** → **Apprentissage de connaissances utiles à la décision**
 - Modélisation du comportement habituel
 - Décision sur la situation par comparaison au profil
- Approche** → **Extraction non supervisée de motifs temporels, multidimensionnels, hétérogènes**



Exemple d'instances d'un même motif (données simulées)

Caractéristiques du problème

- Séquences analysées** : mixtes
- Méthode utilisée** : non supervisée
 - Spécificités individuelles
 - Manque de connaissances *a priori*
- Motifs extraits** : bruités
 - Variabilité dans les valeurs
 - Présence d'interruptions
 - Déformations et translations dans le temps

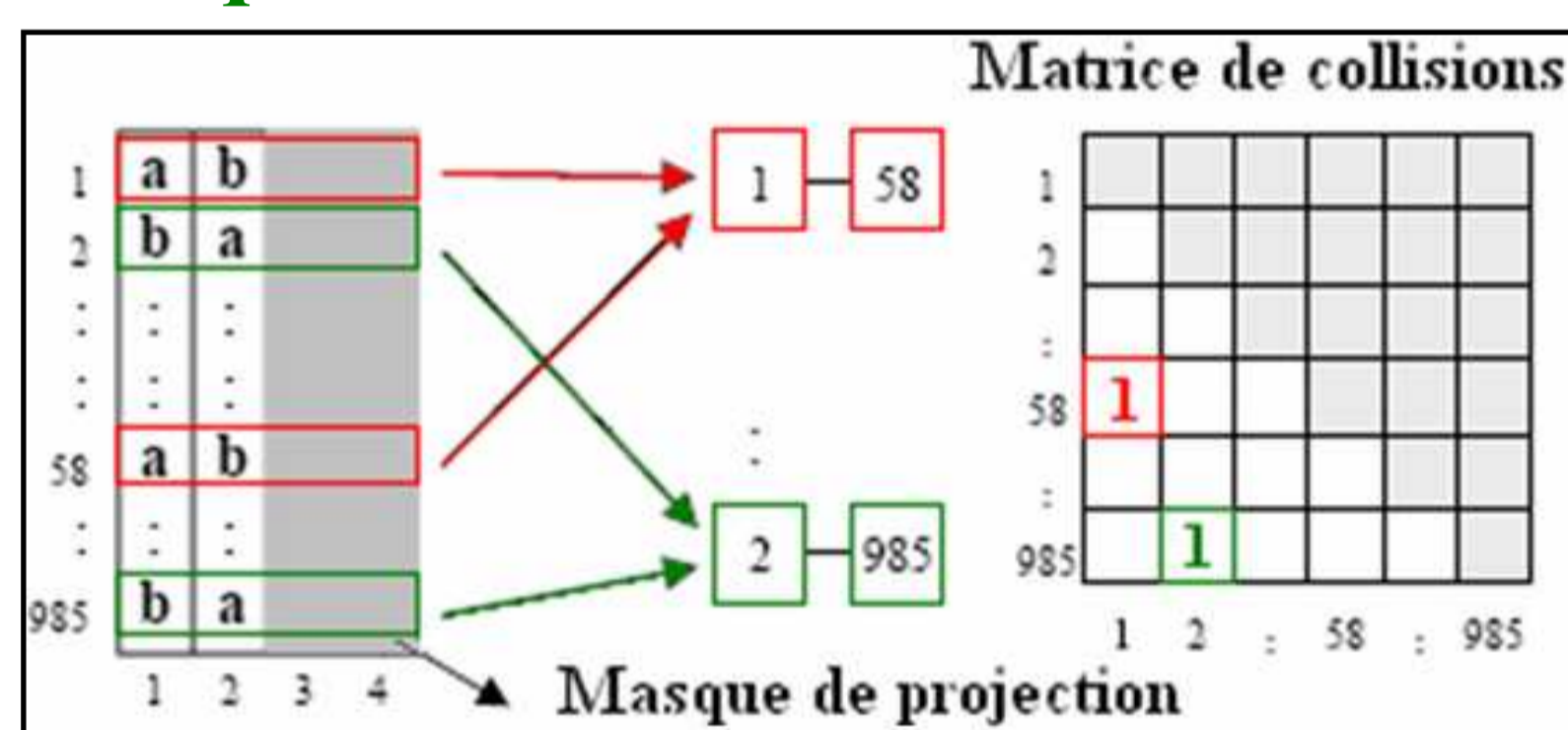
2. Méthode

2.1. Abstraction

- Filtrage moyen pondéré
- Discretisation sur l'axe des valeurs
- Agrégation sur l'axe du temps

2.2. Fouille de caractères – Critères de sélection

1. Fréquence

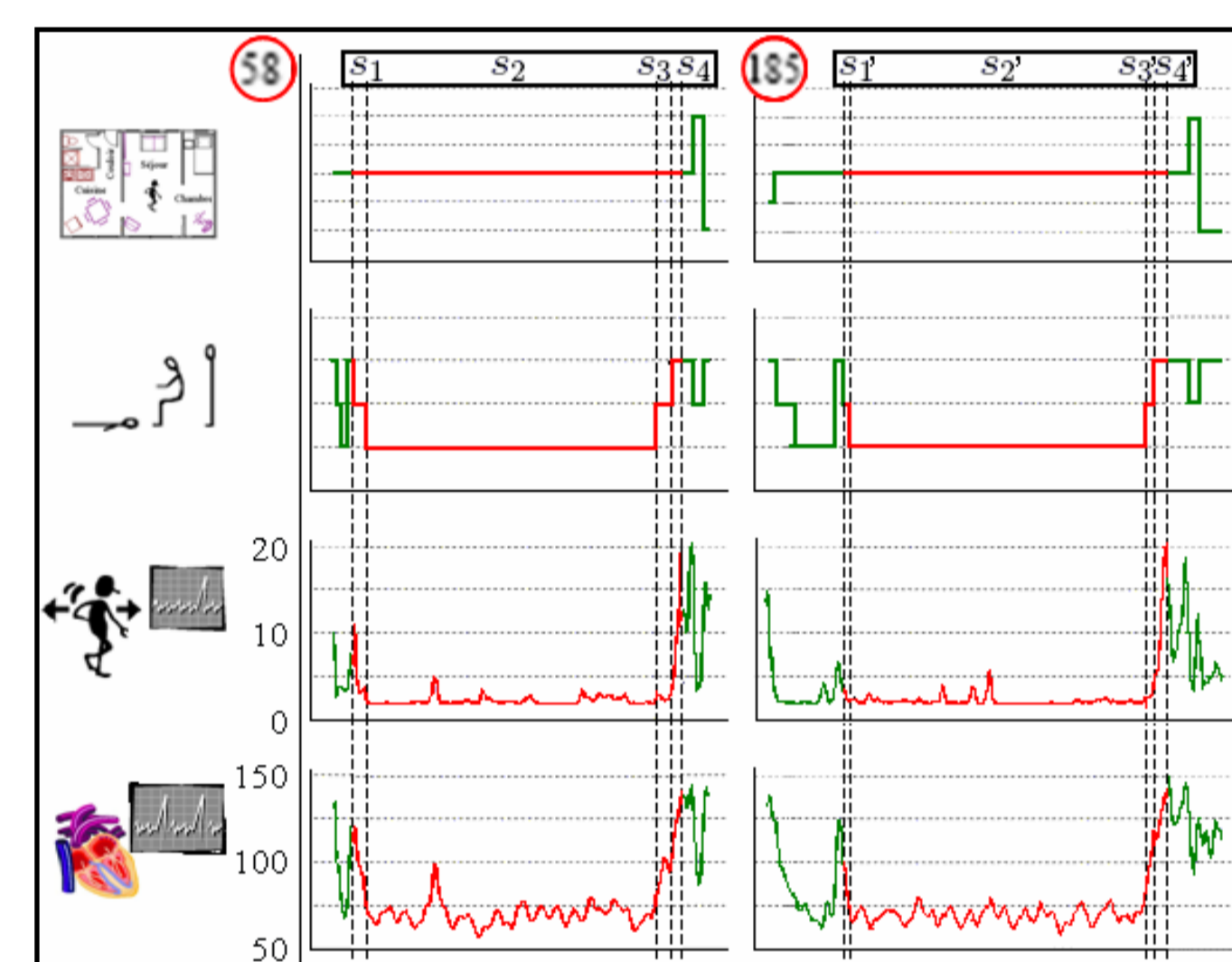


- Projections aléatoires^{1,2} de séquences de base d'un nombre fixe de symboles
- Extension au cas multidimensionnel

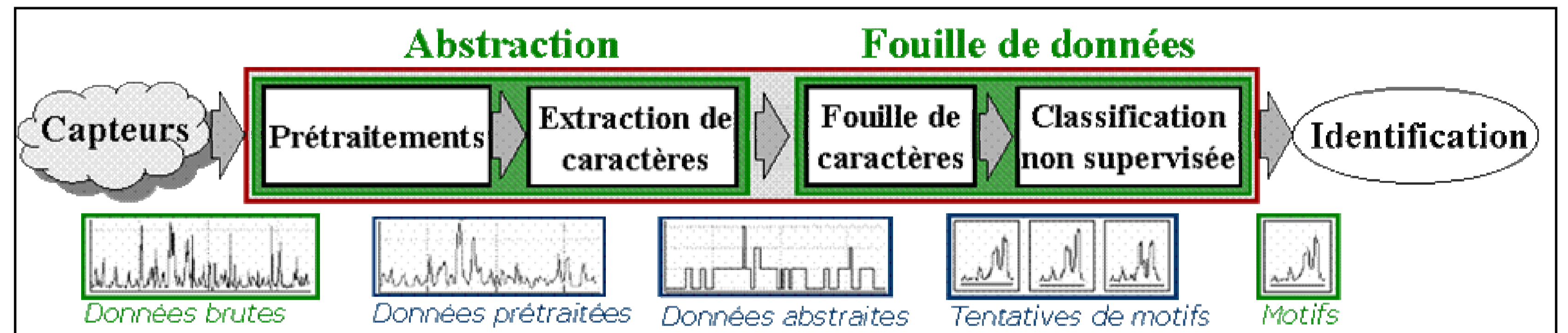
3. Non Redondance

- Groupes des séquences représentatives d'une même instance de motif
- Séquences représentant ces groupes d'instances toutes disjointes

2. Signification



- Extension des récurrences de base
- Critères de collisions et de distance



2.3. Classification non supervisée

- Classification ascendante hiérarchique, selon un seuil de distance

2.4. Distance basée sur la plus longue sous-séquence commune

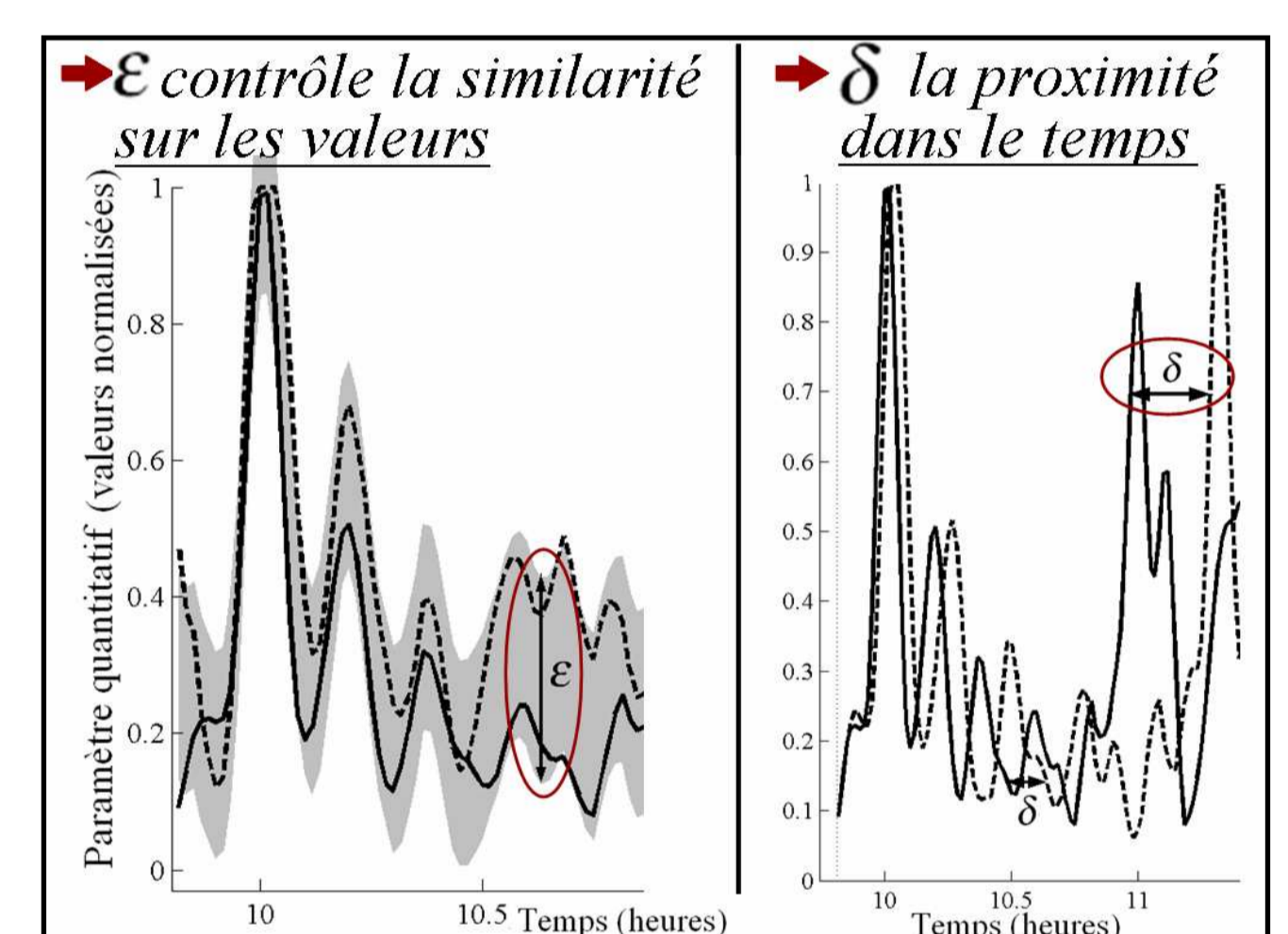
- Mesure non métrique
- Nombres de vecteurs similaires selon deux seuils de similarité³

$$LCSS_{\delta, \epsilon}(A, B)$$

- Distance dans $[0, 1]$ ³:

$$D_{\delta, \epsilon}(A, B) = 1 - \frac{LCSS_{\delta, \epsilon}(A, B)}{\min(n, m)}$$

- Extension au cas hétérogène

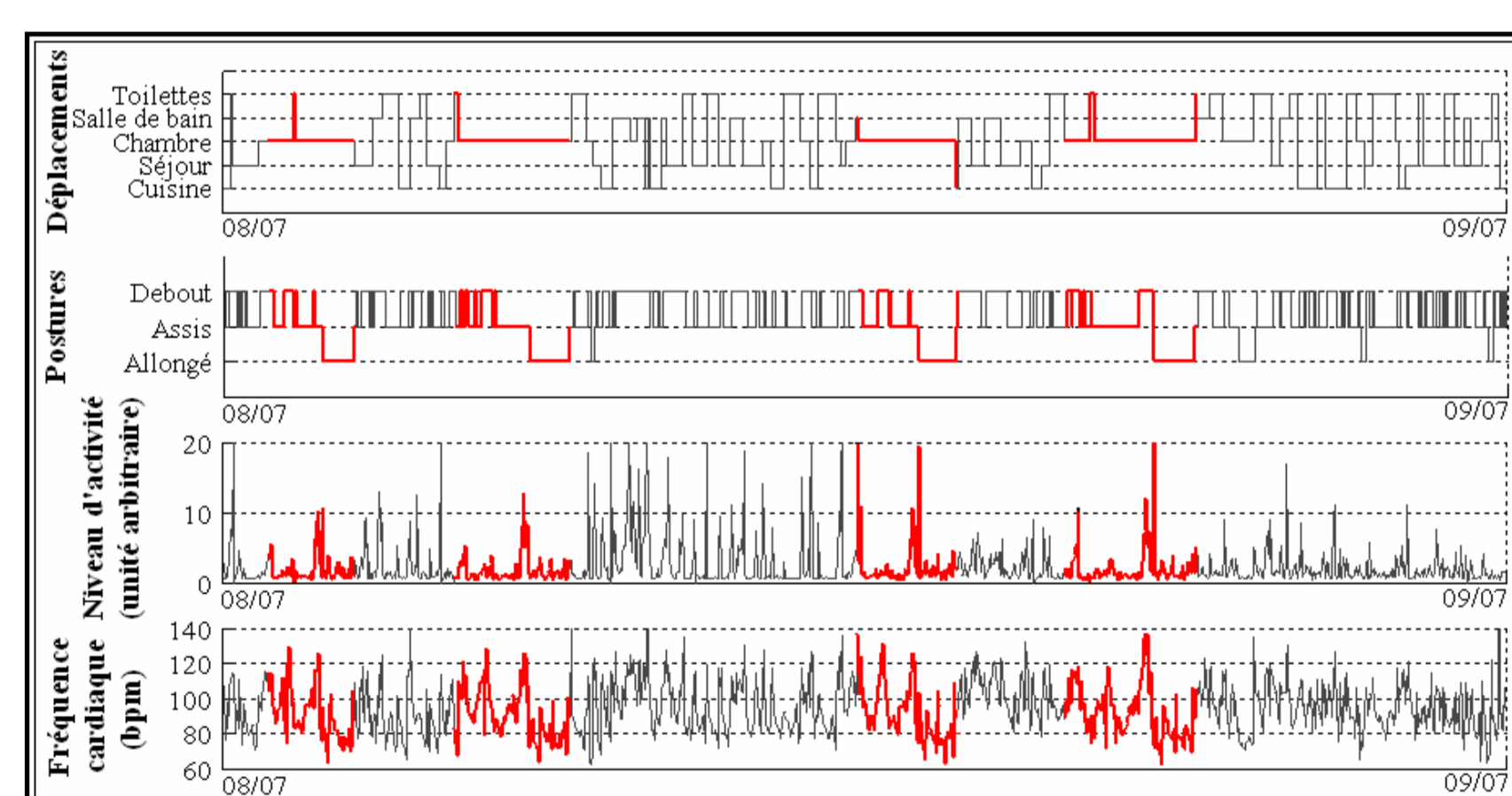


3. Résultats

Sensibilité

- Toutes les instances d'un motif « normalement » modifiées

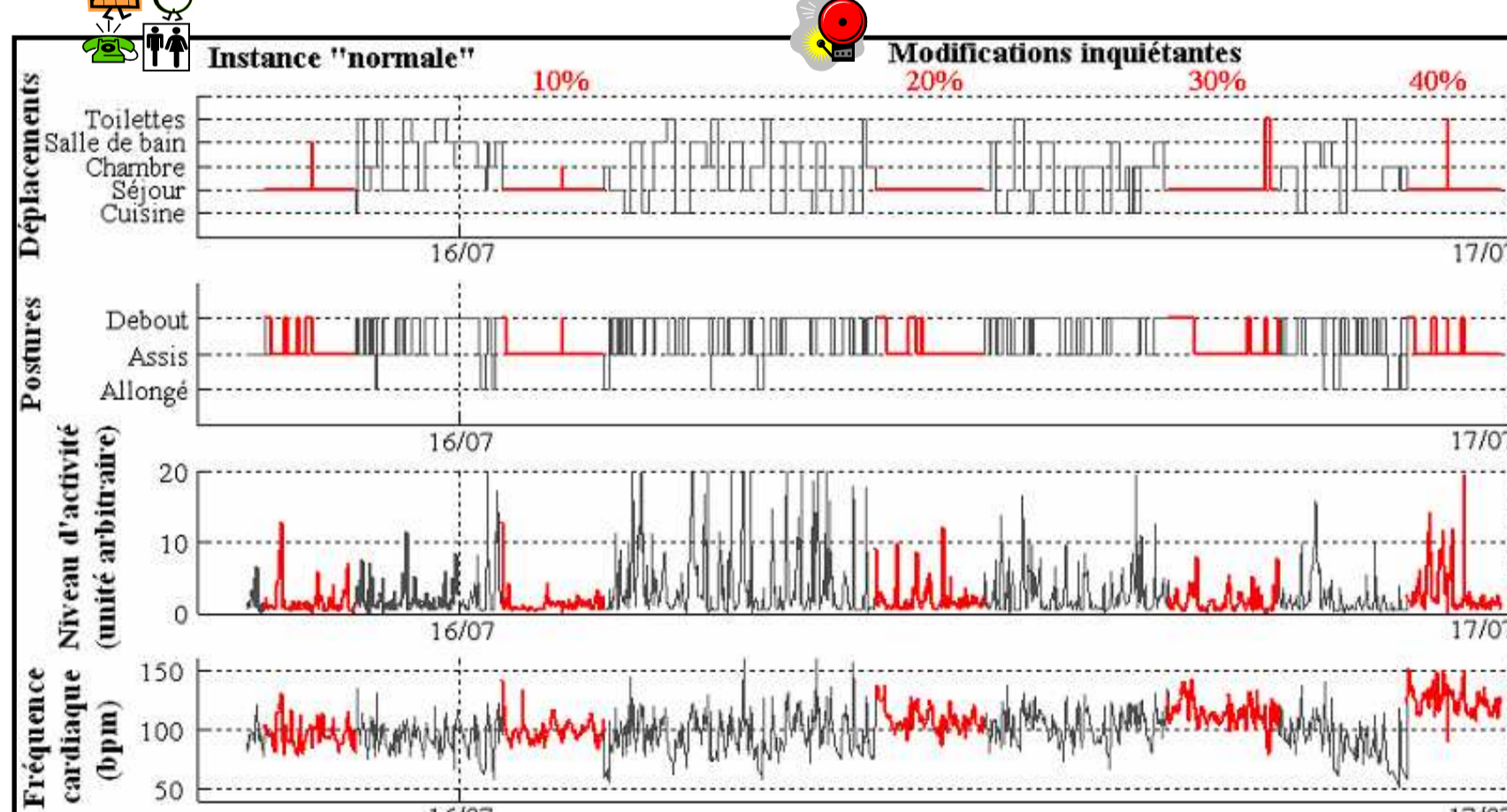
Indices	Identification		Classification	
	S_e	S_p	S_e	S_p
Moyenne	0.71	0.92	0.66	0.79
Écart-type	0.18	0.07	0.34	0.26
Indices parfaits	-	-	35%	60%
Classification parfaite	-	-	20%	-



Spécificité

- Aucune instance d'un motif modifiée de façon inquiétante

Modifications	Normales		Inquiétantes		
	Taux de modification	Taux de reconnaissance	0%	10%	5%
Taux de reconnaissance	77.5%	25%	10%	10%	5%



Motifs extraits de données simulées pour une personne à domicile

Classes	n°1	n°2	n°3	n°4	n°5	n°6
Représentant						
Déplacements						
Postures						
Niveau d'activité						
Fréquence cardiaque						
Effectif	3	5	5	3	10	5
Durée	7h30	3h30	1h15	3h	2h	2h30
Instant d'occurrence	vers minuit	vers 9h	vers 13h ou 20h	vers 15h	entre 16h et 19h	vers 14h ou 21h
Fréquence moyenne recourue	tous les deux jours	presque tous les matins	un jour sur deux	un jour sur deux	une à deux fois par jour	un jour sur deux
Interprétation de l'activité	Nuit de sommeil	Toilette du matin, Petite activité dans la cuisine et repas	Petite période calme au salon	Longue période calme au salon suivie d'une activité plus intense	Repas rapide ou collation suivie d'une activité calme	Période courte de repos

4. Discussion et Conclusion

- Complexité** de la mise en pratique
- Potentialités** de la méthode montrées dans une configuration par défaut et à partir de données simulées
- Validation nécessaire de l'approche sur des **données réelles**

Références

- Buhler J. & Tompa M. (2002). Finding motifs using random projections. *Journal of Computational Biology*, 9(2), 225-242.
- Chiu B., Keogh E. & Lonardi S. (2003), Probabilistic Discovery of Time Series Motifs, *Proc. of the 9th ACM SIGKDD*, Washington, USA, 493-498.
- M. Vlachos, G. Kollios, and G. Gunopulos (2002), Discovering Similar Multidimensional Trajectories, *Proc. of the 18th ICDE*, San Jose, CA, 673-684.