

Analyse spatio-temporelle de systèmes réels décrits par des données issues d'un réseau de capteurs géo-référencés. Application au trafic routier urbain

Laboratoire MAS, Ecole Centrale Paris - Laboratoire LAMSADE, Université Paris Dauphine
Laboratoire GRECIA, INRETS Arcueil - Groupe I3, Département I&R, ENST

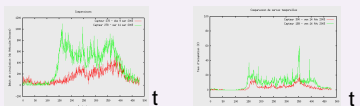
Olivier Carles, Florian De Vuyst, Georges Hébrail, Bernard Hugueney, Marc Joliveau, Geneviève Jomier, Maude Manouvrier, Yosr Naija, Raphaël Rispoli, Gérard Scémama, Laurent Steffan

Données initiales

- Données spatio-temporelles multivariées à précision et fiabilité limitées
- Réseau de capteurs géo-référencés
- Annotations
- Un état de congestion 0/1

Application : réseau urbain

- 1000+ capteurs
- Echantillonnage : 3 min
- 1000+ jours



Débit et taux d'occupation à un capteur (2 instances)

Travaux en cours

Enrichissement de l'entrepôt par résumés, SIG, interaction utilisateur...

Approche symbolique spatio-temporelle
Prototype

Simulation, prévision

Représentations numériques-symboliques
Réduction dimensionnelle très concluante

Mise en ligne de l'entrepôt
Adaptation de l'entrepôt par apprentissage

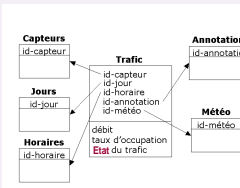
Conclusions et perspectives

Publications

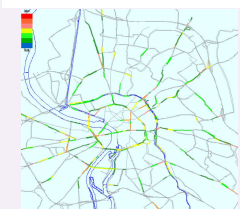
- C. Bauder-Medeiros, O. Carles, F. De Vuyst, G. Hébrail, B. Hugueney, M. Joliveau, G. Jomier, M. Manouvrier, Y. Naija, G. Scémama, L. Steffan, Vers un entrepôt de données pour le trafic routier, Proc. Conférences EDA '06, Versailles (2006)
- Bernard Hugueney, Adaptive Segmentation-Based Symbolic Representations of Time Series for Better Modeling and Lower Bounding Distance Measures, 10th European Conference on Principles and Practice of Knowledge Discovery in Databases PKDD, Septembre 2006, Berlin, Allemagne.
- B. Hugueney, G. Hébrail, Y. Lechevallier, Computing Summaries of Time Series Databases with Clustering and Segmentation, 10th ICDM Conference, Data Science and Classification Juillet 2006, Slovénie.
- B. Hugueney, Représentations symboliques adaptatives de séries temporelles : principes et algorithmes de construction, EGC 2006.
- M. Joliveau, F. De Vuyst, Recherche de motifs dans les séries temporelles de trafic routier urbain. Définition d'une architecture d'analyse, Rapport de recherche MAS n°0605, octobre 2006.

Entrepôt de données

Schéma de données



Exploration visuelle dynamique



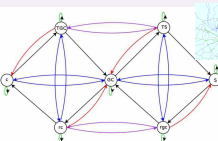
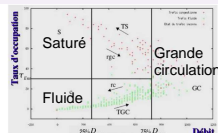
Attributs des dimensions d'analyse

Capteurs	Jours
Coordonnées géographique	Date
Selon la topologie du réseau	Nom de jour
Axe, zone,...	Calendrier
Selon les découpages administratifs	Semaine, mois, année
Quartier, commune, agglomération	Activité scolaire
Selon le type d'activité	Vacances : Noël, Février,...
Zone industrielle, zone commerciale,	Périodes scolaires
Zone de loisir, zone d'habitation	Veille / lendemain de vacances
Zone d'activité scolaire, universitaire	Ponts
	Activité professionnelle
	Férié, ouvrable
	Marchés
	Week-end, semaine
	Pont
	Activité socioculturelle
	Manifestations, festivals

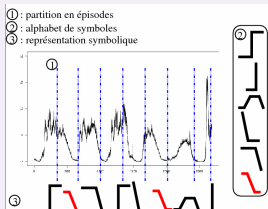
Exemple d'interrogation

Quels sont les axes fluides dans les zones commerciales l'après-midi des veilles de vacances de Noël ?

Représentations symboliques



Génération de 7 états selon un diagramme de transition
→ intelligibilité

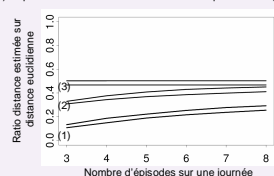


Extrait de représentation symbolique sur une semaine (débit)

Découpage en épisodes
→ réduction de dimensionnalité
Symboles : étiquettes de classes de comportements sur les épisodes considérés
→ intelligibilité

Qualité de l'approximation de la distance euclidienne réelle

- Représentations purement symboliques (classique et adaptative)
- Représentations avec informations numériques globales (idem)
- Représentations avec informations numériques locales (idem)



Représentations hybrides numériques et symboliques :
L'information numérique permet des calculs plus précis (distances)
→ compromis encombrement/précision

Analyse spatio-temporelle

- En entrée : X
- Tous les capteurs
 - Tous les jours ouvrés
 - Données débit

Réduction : $I^T = J^T = O(10)$.

1. Projecteur spatial π_x :

$$\pi_x \mathbf{X} = \mathbf{P}\mathbf{P}^T \mathbf{X}, \mathbf{R}_x = \mathbf{P}^T \mathbf{X} \text{ de taille } I^T \times J.$$

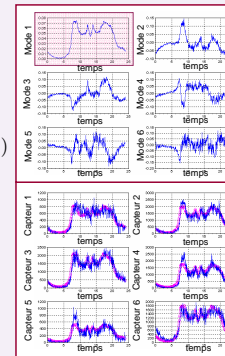
2. Projecteur temporel π_t :

$$\pi_t \mathbf{X} = \mathbf{X}\mathbf{Q}\mathbf{Q}^T, \mathbf{R}_t = \mathbf{X}\mathbf{Q} \text{ de taille } I \times J^T$$

3. Projecteur spatio-temporel $\pi_{x,t} = \pi_x \pi_t = \pi_t \pi_x$:

$$\pi_{x,t} \mathbf{X} = \mathbf{P}\mathbf{R}\mathbf{Q}^T, \mathbf{R}_{x,t} = \mathbf{P}^T \mathbf{X}\mathbf{Q} \text{ de taille } I^T \times J^T.$$

Analyse en Composantes Principales temporelles



Résultats :
- Les 6 premiers modes temporels (ACP-t) en colonne définissent la matrice Q (J=6)

Débits, 1 jour ouvré, 6 capteurs
- Débit réel
- Projection temporelle sur le 1er mode

→ 95 % de l'information restituée avec le 1er mode