

Réflexions à partir d'expérimentations de données téléphoniques et de données de réseaux sociaux

Bernard Elissalde, Université de Rouen, UMR IDEES

Julie Fen-Chong, Université de Bourgogne Franche Comté - UMR ThéMA

Françoise Lucchini, Université de Rouen, UMR IDEES

Journée d'étude **Données individuelles localisées à l'heure du RGPD**
Axe INFormation TERritoriale locale INFTER

Vendredi 15 novembre 2019 | Paris Campus Paris Diderot

Les données géolocalisées de téléphonie mobile et de réseaux sociaux

Chercheurs

Sciences de la complexité

Nouvelles sources sur la mobilité
Human mobility, models and applications, Barbosa et al. 2018

- 1) Accès aux données
- 2) Forces et limites de ces données
- 3) Réflexions méthodologiques

Usagers

Producteurs de données :

- Passifs
- Volontaires

Mais n'ont pas conscience de leur utilisation

Téléphonie 3-4 millions d'utilisateurs par jour
en IDF ~ 60-100 millions événements
Twitter 247 000 utilisateurs ~ 1,2 m. tweets
sur une année (2015)

Contexte d'accès aux données

- Thèse 2009-2012.

Contrat CIFRE entre les partenaires signé en décembre 2010, accès aux données d'Orange à l'automne 2010

- Depuis 2008 « Pulsations urbaines, usage éthique des traces de téléphonie mobile pour la recherche en géographie urbaine »

PPP avec la délégation d'Orange en Normandie

- 2019-2015 « Modélisation des usages des territoires et des mobilités induites. Axe Pulsations urbaines ».

Sonde de captage Twitter

- 2021-2018 « Modélisation des trajectoires des mobilités terrestres des passagers maritimes »

Présentation des données

Information	Données Téléphonie	Données Twitter
Entité	Individu (IMSI/n° alias)	Individu (ID Usager)
Granularité temporelle	Seconde	Seconde
Granularité spatiale	Antenne (agrégat spatial)	GPS (point)

Présentation des données

Avec Flux vision

Information	Données téléphonie
Entité	Groupe de 20 individus identifiés par similarité spatio-temporelle
Granularité temporelle	Plusieurs heures
Granularité spatiale	Antennes ou communes

Les données de téléphonie mobile : des données passives

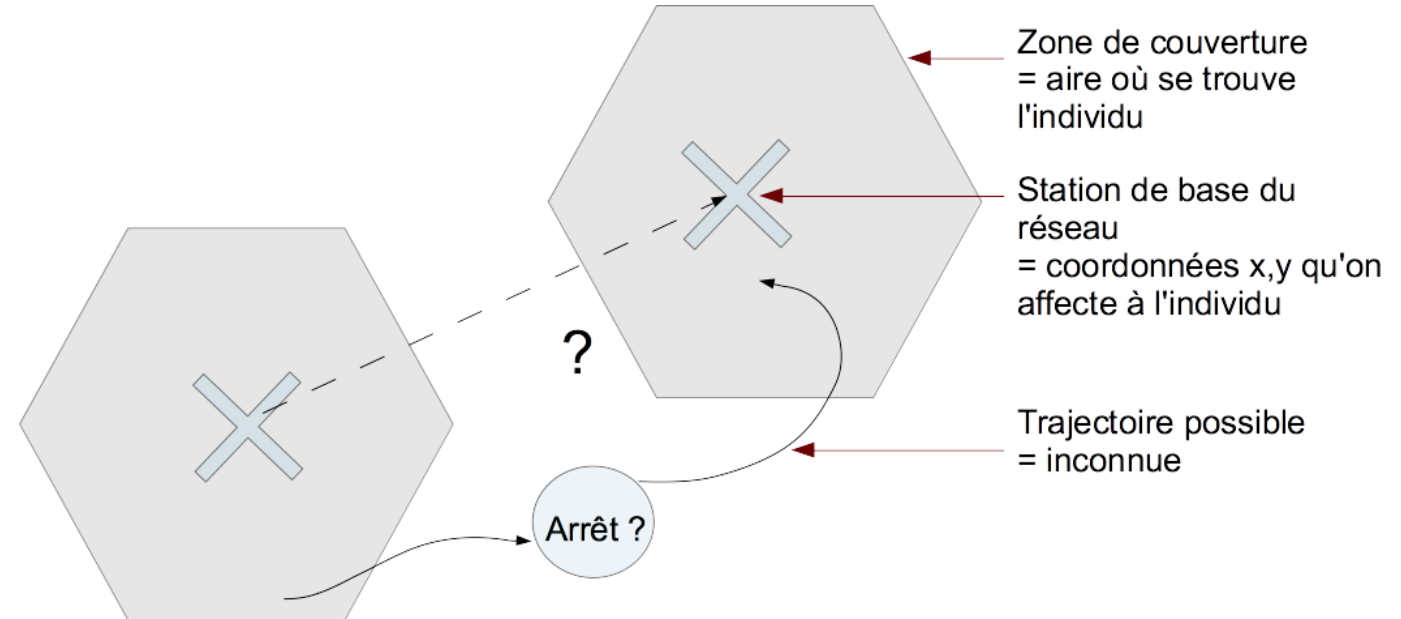
Etanchéité des données

Données « de facturation »

appel reçus & émis avec données sur l'utilisateur, son adresse etc.

Données techniques

Tous les événements du réseau mais sans information sur l'utilisateur



Imperfections pour la recherche en géographie

Précision des données

- données manquantes
- granularité spatiale et temporelle évolutive
- représentativité
- anonymat

Sémantique

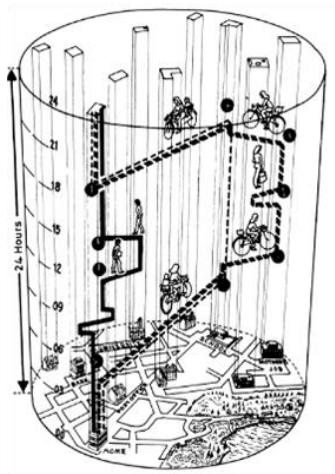
- relation sociologique entre un événement et le déclenchement de l'activité géolocalisée
- uniquement le mouvement, pas origine/destination

Disponibilité

- accès/gratuité des données

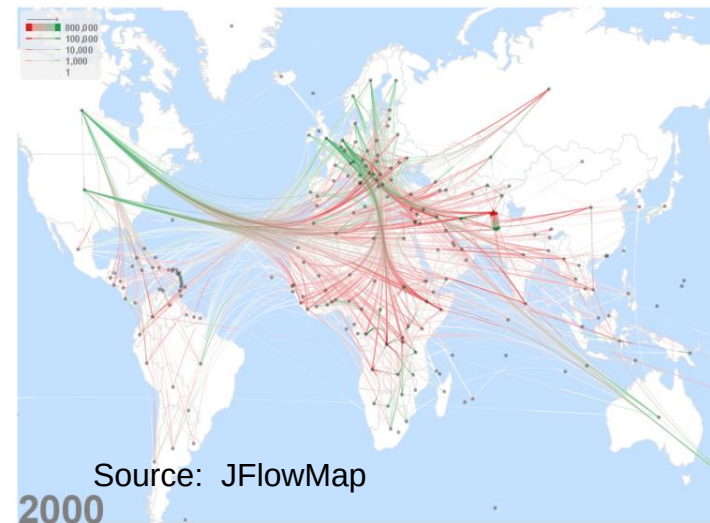
La mobilité : de la variabilité individuelle aux rythmes collectifs

Trajectoire individuelle



LENNTORP B.,
1978

Mouvements collectifs



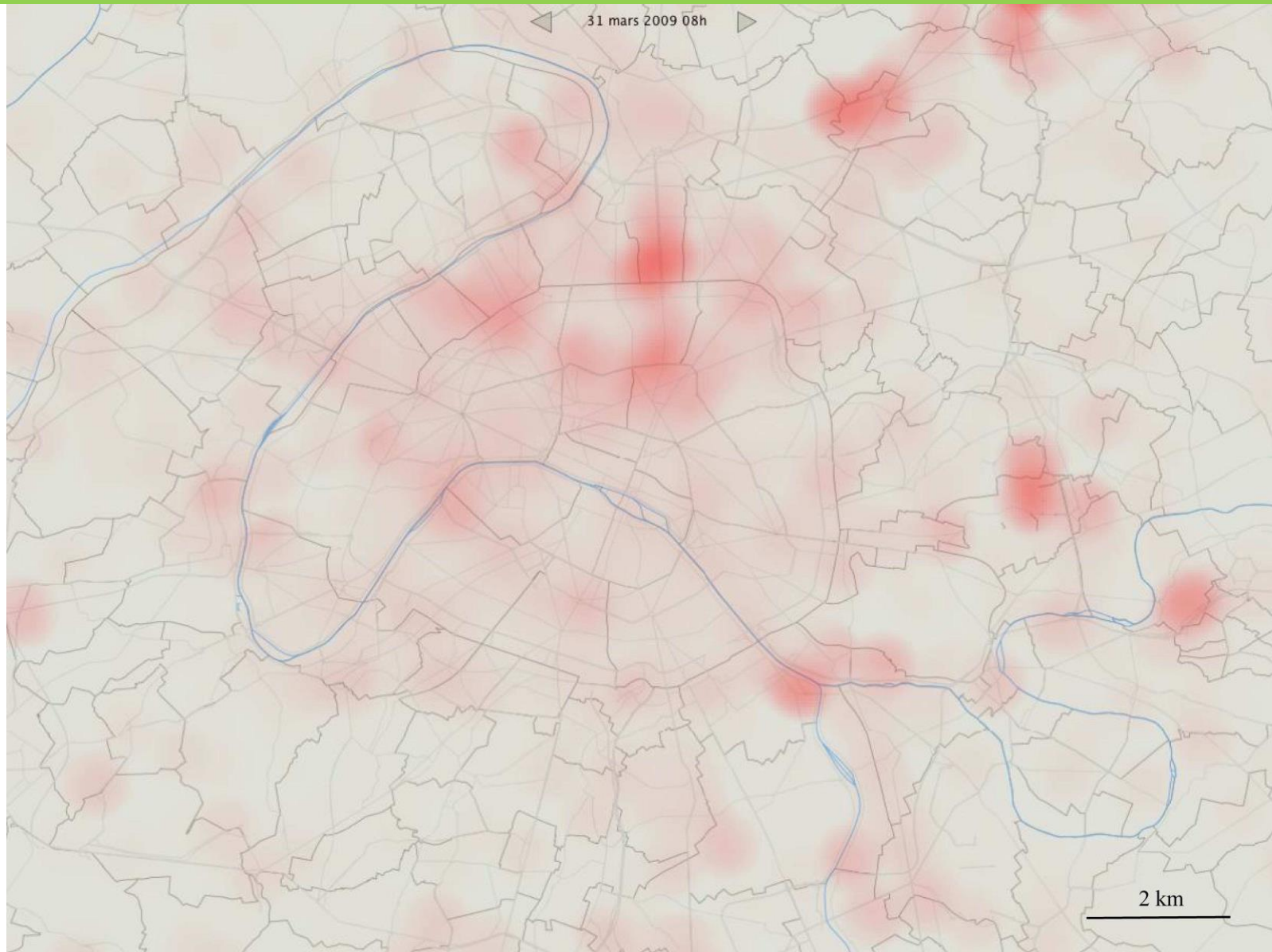
Source: JFlowMap

2000

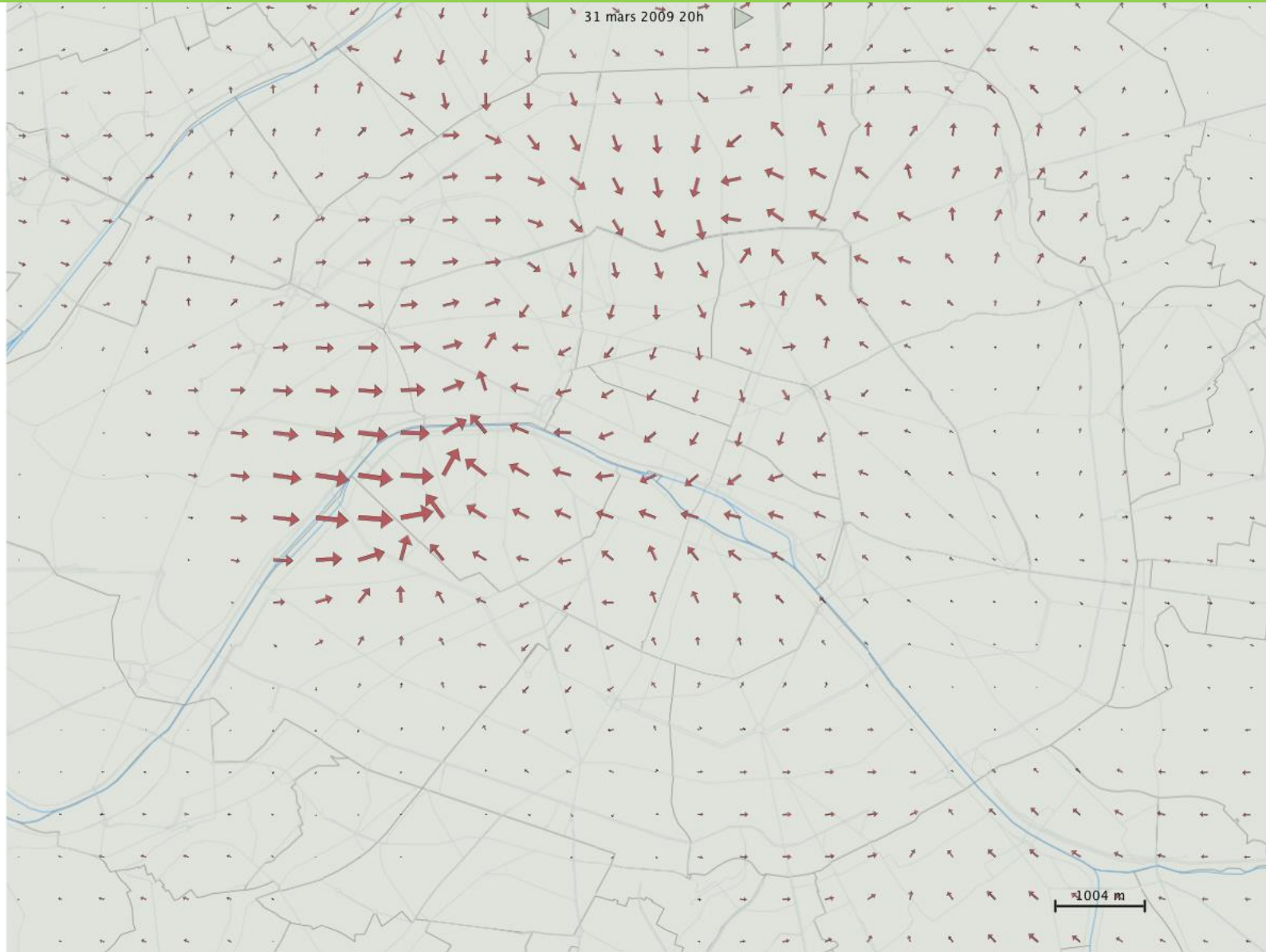
Méthodes pour une connaissance du fonctionnement spatio-temporel

Méthode avec données de téléphonie mobile	Identification du phénomène
Agrégation d'utilisateurs autour d'une antenne	Fréquentation touristique Concentration, polarisation
Vecteurs Moyens à partir des flux	Orientation du mouvement
Agrégation de flux d'une antenne à une autre	Mouvement
Détection de communautés	Communauté de lieux
Graphes	Hubs + trajectoires principales Structure des mobilités
Détection d'événements Comparaison de graphes Indices de similarités spatiales, de similarités hiérarchiques (Top 10)	Anormalité des concentrations $\neq$ routine (sommets, arêtes)

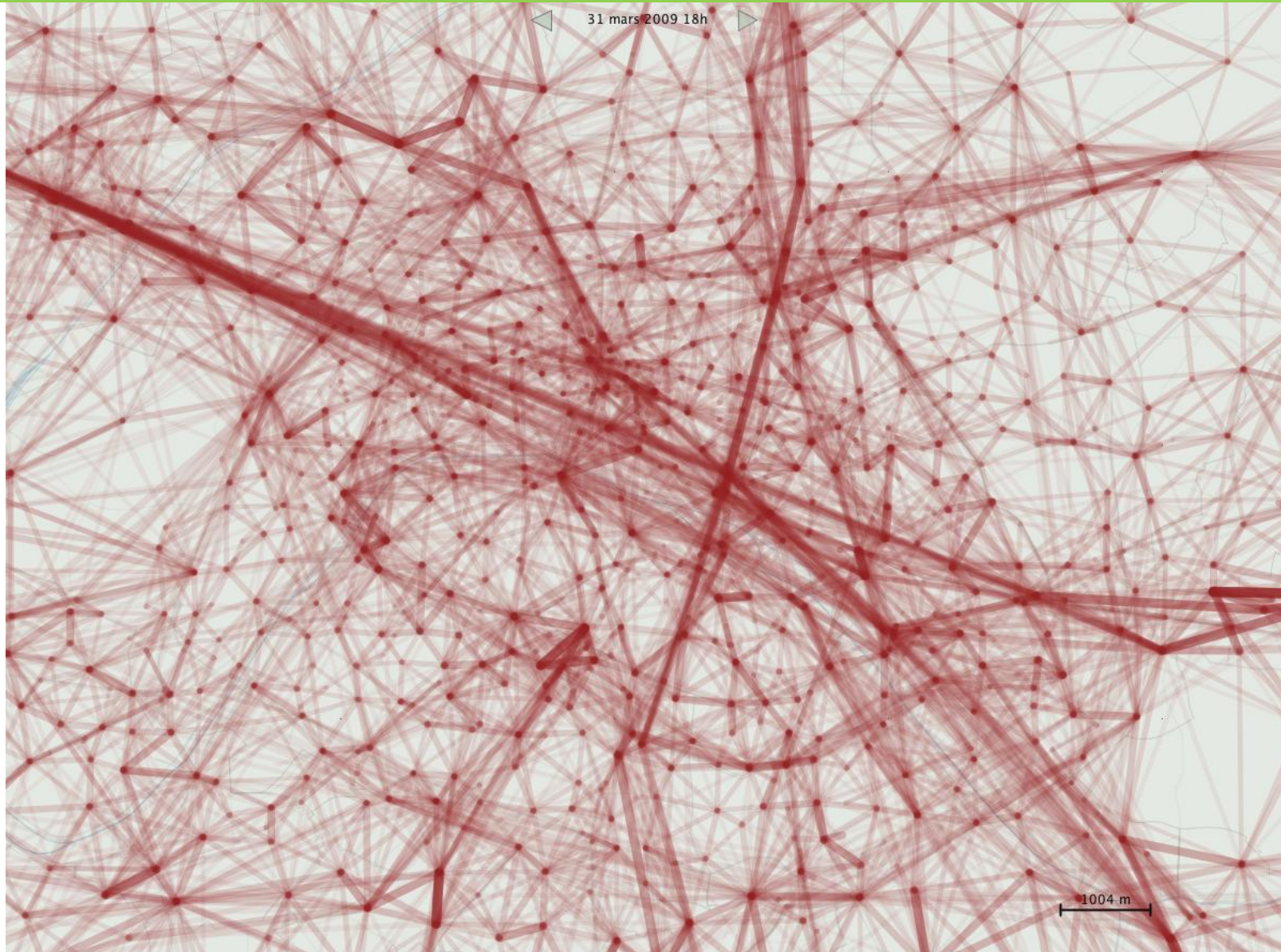
Concentration de population



Orientation du mouvement



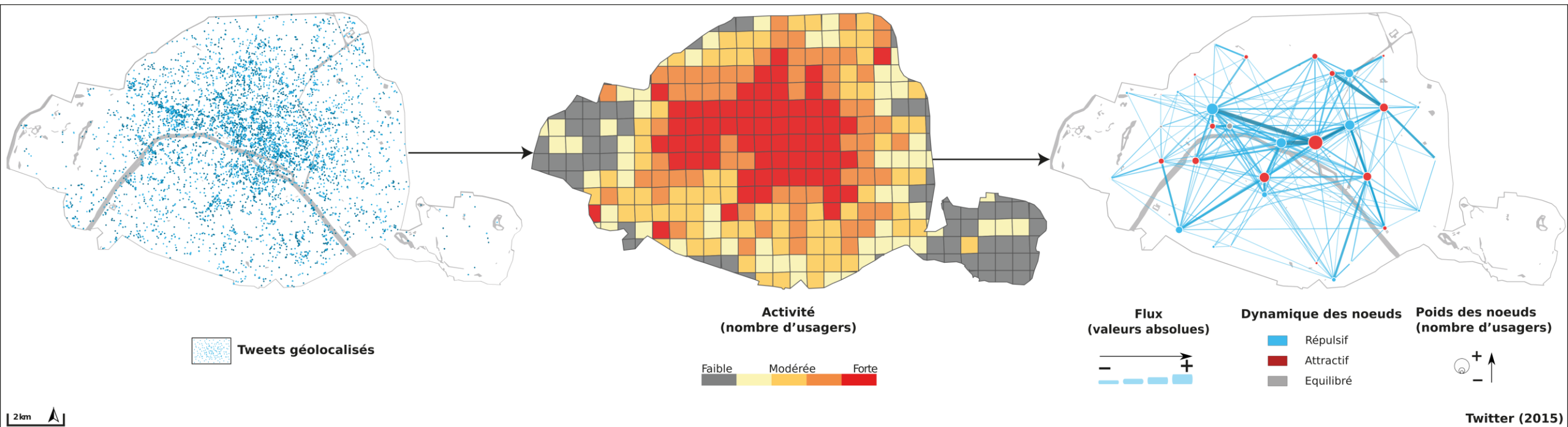
Mouvement entre antennes



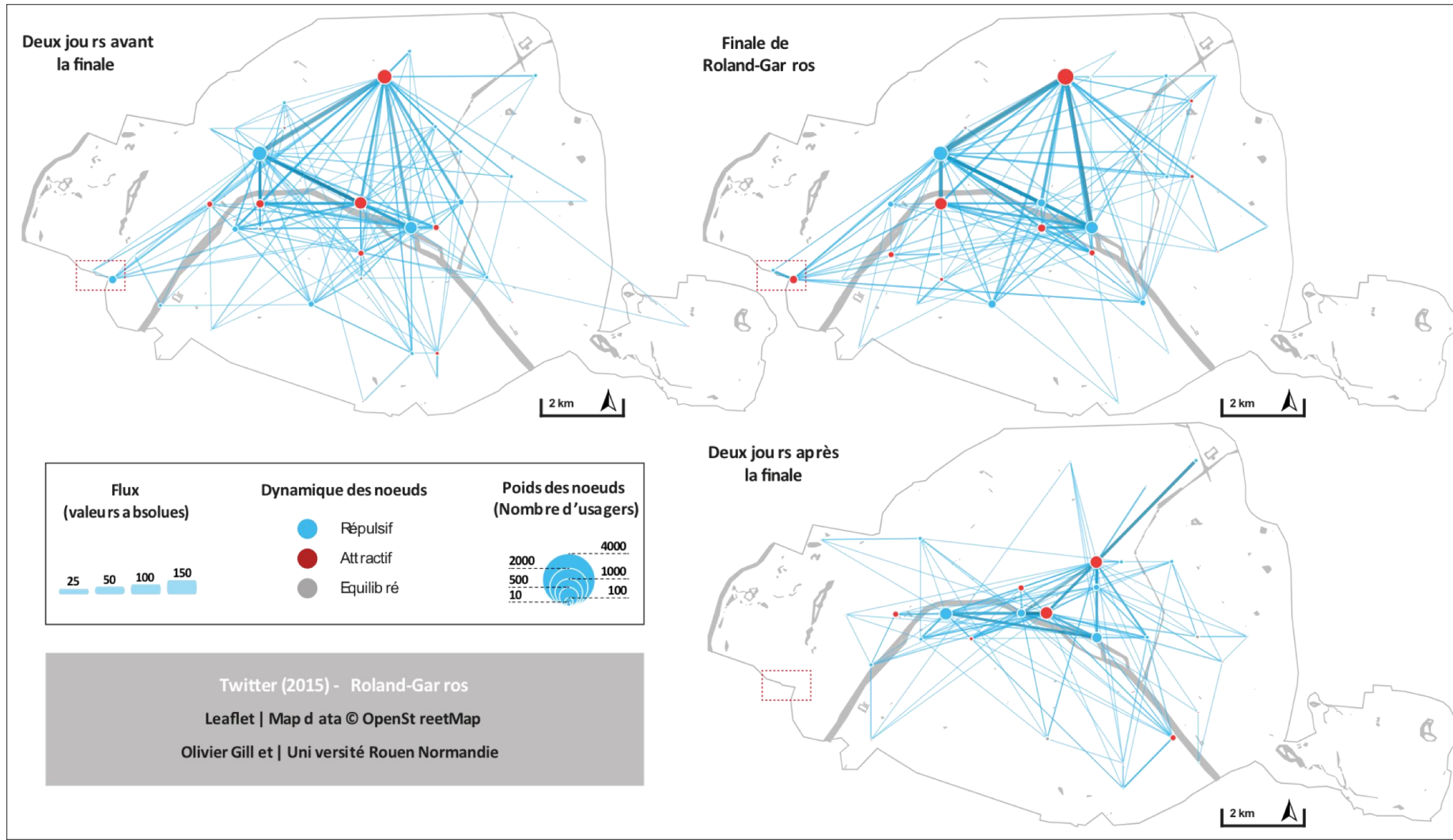
Méthodes pour une connaissance du fonctionnement spatio-temporel

Méthodes avec Twitter	Identification du phénomène
Agrégation surfacique Indice de surreprésentation LISA	Concentration + ou – écarts types / moyenne
Graphe par clustering d'agrégation spatiale et temporelle - Proximité spatiale - Nb min points - Dynamique (attractif/répulsif/neutre) - Activité (nb usagers) DBScan, MobilityGraph, STDBScan	Mouvements Hubs, arêtes
Comparaison des graphes	Arythmie, détection d'évènement

Concentrations urbaines et mouvements à partir de Twitter



Détection d'évènements à partir de Twitter



Construire de nouveaux profils d'utilisateurs



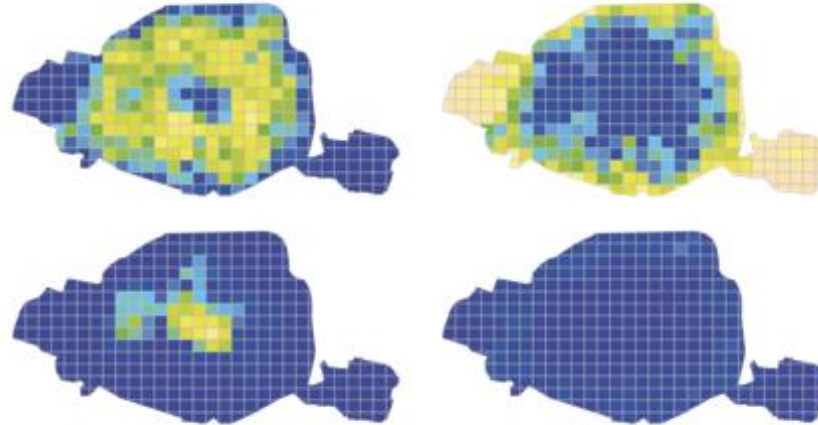
Not very chatty
and mobile



Very chatty,
not mobile



Chatty and
very mobiles



Parameters

User activity
User mobility
Maximal distance
Spatial attractiveness
Preferential sector

Bilan Ce qu'on apprend de la ville

Représenter la ville dans sa fonction de lieu d'échange et non pas l'individu

- Un lieu d'échanges et d'interaction sociale
- Pulsations de la ville : rythmes routiniers habituels et arythmies
- Mobilités et déplacements préférentiels
- Découpages spatio-temporels
- Attractivité des quartiers
- Quels usages de la ville/ grande variabilité des concentrations d'utilisateurs

Merci de votre attention

Journée d'étude **Données individuelles localisées à l'heure
du RGPD** Axe INformation TERritoriale locale INFTER

Vendredi 15 novembre 2019 Paris Campus Paris Diderot