

# 実験・検証の基本的な考え方(1)

位置情報を利活用するに当たって、個人が特定される主なリスクとして以下のものが挙げられている。（位置情報プライバシーレポートより）

## 【位置情報を取り扱うに際しての個人が特定されるリスクの例】

- ・ 位置情報に伴う氏名や属性情報等の記述から個人が特定されてしまう。
- ・ 詳細な位置情報と時間から個人が特定されてしまう。
- ・ 位置情報と他のデータが、位置情報を用いてマッチングされ、その結果個人が特定されてしまう。
- ・ 特徴的な位置情報の履歴から生活圏や行動パターンがわかり、個人が特定されてしまう。

## 実験・検証の基本的な考え方(2)

位置情報を利活用するに当たって想定される主なリスクへの対策となる加工方法の例として、以下のものが挙げられている。(位置情報プライバシーレポートより)

### 【有効と考えられる加工方法の例】

- ・ 直接あるいは組み合わせで個人が特定できる情報の削除、仮名化
- ・ 組み合わせで個人が特定できる情報の一般化、ランダム化
- ・ 位置情報(時間)のより広いエリア(時間帯)への一般化、違う位置(時間)へのランダムな置き換え(図表1)
- ・ 生活圏情報や行動パターンの削除、一般化、置き換え(図表2)
- ・ 仮名の短い時間での更新、長い履歴の削除、分割や間引き(図表3)
- ・ 位置情報の取得間隔の適切な設定(極端に短い間隔にしない)
- ・ 上記の手法を用い位置情報と属性情報を適切に加工し、全ての属性に対して、同じ位置情報(移動の軌跡を含む)が複数ある状況を作り出す。(図表4)

## 実証・検証における検証内容

---

位置情報プライバシーレポートでは「十分な匿名化」された情報は、契約約款等の同意で第三者に情報提供が可能とされている。

もともと、「十分な匿名化」の水準は、あらかじめ一律に示すことは困難とされており、データセットの性質やその利活用の態様に応じて異なる性質のものであるため、実証・検証においては前頁に記載する「有効と考えられる加工方法の例(非特定化・ランダム化・集計化・秘匿化、等)」を複数用いることがリスク対策として適切なものであることを確認すると共に、更に安心安全な加工方法を探るものとする。

なお、具体的な加工水準を一律に示すことが困難である以上、具体的な加工方法は実証・検証を行う事業者に委ねることが適切と考える。

また、同レポートにおいては利用される情報が通信の場所、日時及び利用者・端末識別符号に限定されているが、年齢、性別等の個人情報を含めた場合であっても、十分な匿名化を行うための加工方法例として適切なものであることも併せて検証する。

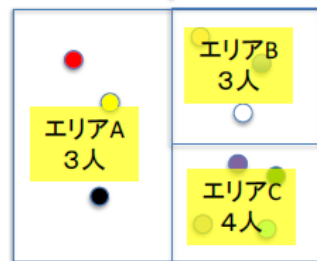
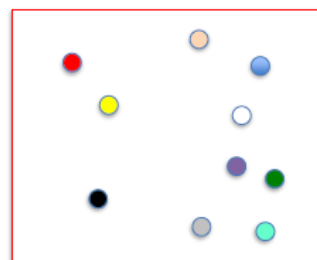
# 図表1: 位置情報の一般化、ランダム化の例

(10名のある時刻における位置情報を扱うケース)

| 位置   |
|------|
| 位置1  |
| 位置2  |
| 位置3  |
| 位置4  |
| 位置5  |
| 位置6  |
| 位置7  |
| 位置8  |
| 位置9  |
| 位置10 |



| 位置   |
|------|
| エリアA |
| エリアA |
| エリアA |
| エリアB |
| エリアB |
| エリアB |
| エリアC |
| エリアC |
| エリアC |
| エリアC |



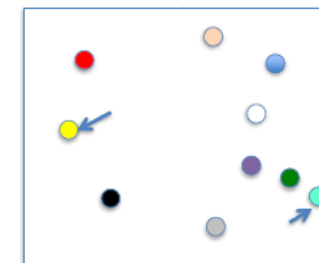
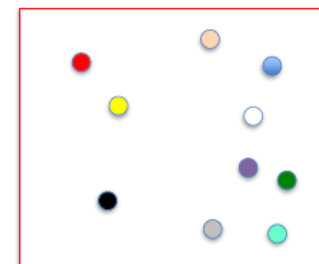
※位置を適切なサイズのエリアに拡大させて、エリアごとに十分な人数を確保する。

(10名のある時刻における位置情報を扱うケース)

| 位置   |
|------|
| 位置1  |
| 位置2  |
| 位置3  |
| 位置4  |
| 位置5  |
| 位置6  |
| 位置7  |
| 位置8  |
| 位置9  |
| 位置10 |



| 位置    |
|-------|
| 位置1   |
| 位置2'  |
| 位置3   |
| 位置4   |
| 位置5   |
| 位置6   |
| 位置7   |
| 位置8   |
| 位置9   |
| 位置10' |

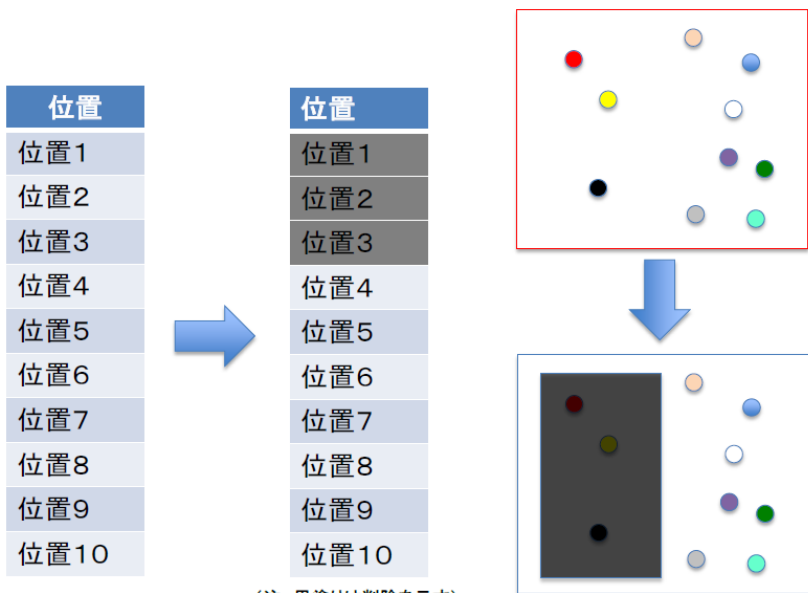


※いくつかの位置の場所をランダム化（不規則にずらす、置き換える）することにより知られたくない位置が確定的にもれることを防ぐ。

出典: 位置情報プライバシーレポートより

# 図表2:生活圏情報等の削除の例

(10名のある時刻における位置情報を扱うケース)



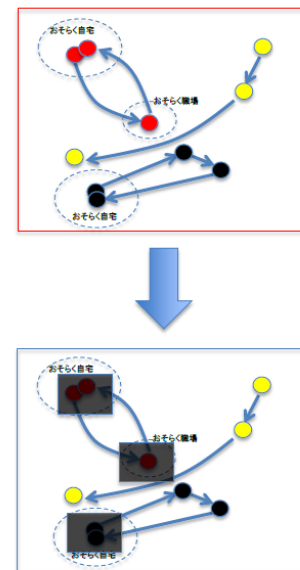
(注. 黒塗りは削除を示す)

※ 自宅の位置情報を削除。  
必要に応じて職場や病院などの位置情報を削除(分析目的による)

(3名の移動履歴を扱うケース)

| ID   | 時刻    | 位置  |
|------|-------|-----|
| ID01 | 08:00 | 位置1 |
| ID01 | 09:00 | 位置2 |
| ID01 | 20:00 | 位置1 |
| ID02 | 08:00 | 位置3 |
| ID02 | 10:00 | 位置4 |
| ID02 | 15:00 | 位置5 |
| ID02 | 21:00 | 位置3 |
| ID03 | 08:00 | 位置6 |
| ID03 | 09:00 | 位置7 |
| ID03 | 20:00 | 位置8 |

| ID   | 時刻    | 位置  |
|------|-------|-----|
| ID01 | 08:00 | 位置1 |
| ID01 | 09:00 | 位置2 |
| ID01 | 20:00 | 位置1 |
| ID02 | 08:00 | 位置3 |
| ID02 | 10:00 | 位置4 |
| ID02 | 15:00 | 位置5 |
| ID02 | 21:00 | 位置3 |
| ID03 | 08:00 | 位置6 |
| ID03 | 09:00 | 位置7 |
| ID03 | 20:00 | 位置8 |



※ ID01の自宅と職場、ID02の自宅を削除。

出典: 位置情報プライバシーレポートより

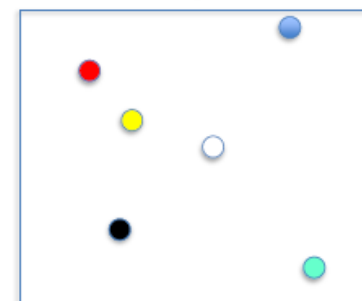
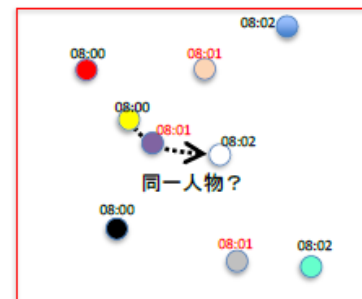
# 図表3: 更新化間隔間引きの例

(複数名の時刻をまたぐ位置情報を扱うケース)

| 時刻    | 位置  |
|-------|-----|
| 08:00 | 位置1 |
| 08:00 | 位置2 |
| 08:00 | 位置3 |
| 08:01 | 位置4 |
| 08:01 | 位置5 |
| 08:01 | 位置6 |
| 08:02 | 位置7 |
| 08:02 | 位置8 |
| 08:02 | 位置9 |



| 時刻    | 位置  |
|-------|-----|
| 08:00 | 位置1 |
| 08:00 | 位置2 |
| 08:00 | 位置3 |
| 08:01 | 位置4 |
| 08:01 | 位置5 |
| 08:01 | 位置6 |
| 08:02 | 位置7 |
| 08:02 | 位置8 |
| 08:02 | 位置9 |



※ 時間的に近接した位置情報を間引いて、移動の軌跡の推定を防止する。

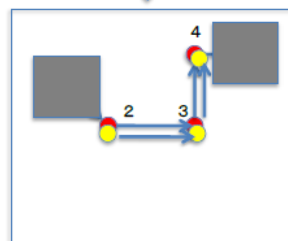
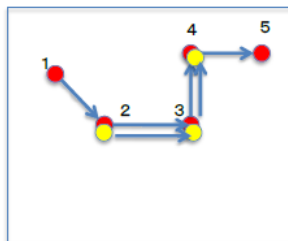
出典: 位置情報プライバシーレポートより

# 図表4: 位置情報の複数化の例

(複数名の移動履歴を扱うケース)

| ID   | 時刻    | 位置  |
|------|-------|-----|
| ID01 | 08:00 | 位置1 |
| ID01 | 09:00 | 位置2 |
| ID02 | 09:00 | 位置2 |
| ID01 | 10:00 | 位置3 |
| ID02 | 10:00 | 位置3 |
| ID01 | 11:00 | 位置4 |
| ID02 | 11:00 | 位置4 |
| ID01 | 12:00 | 位置5 |

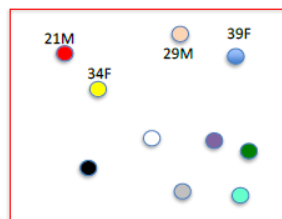
| ID   | 時刻    | 位置  |
|------|-------|-----|
| ID01 | 08:00 | 位置1 |
| ID01 | 09:00 | 位置2 |
| ID02 | 09:00 | 位置2 |
| ID01 | 10:00 | 位置3 |
| ID02 | 10:00 | 位置3 |
| ID01 | 11:00 | 位置4 |
| ID02 | 11:00 | 位置4 |
| ID01 | 12:00 | 位置5 |



※ 同じ行動の人がいる軌跡を使う。  
 《同時間・同位置》の人は現実には存在しないため、位置のまとめ・時間のまとめ、および系列長の制限を行う。  
 同じ行動の人が多数いるほど個人識別が起りにくい。

(10名のある時刻における位置情報と属性情報(年齢・性別)を扱うケース)

| 年齢    | 性別    | 時刻    | 位置    |
|-------|-------|-------|-------|
| 21    | 男     | 08:00 | 位置1   |
| 34    | 女     | 08:00 | 位置2   |
| 29    | 男     | 08:00 | 位置3   |
| 39    | 女     | 08:00 | 位置4   |
| 57    | 男     | 08:00 | 位置5   |
| 49    | 女     | 08:00 | 位置6   |
| 55    | 男     | 08:00 | 位置7   |
| ..... | ..... | ..... | ..... |



| 年齢    | 性別    | 時刻    | 位置    |
|-------|-------|-------|-------|
| 20代   | 男     | 08:00 | エリアA  |
| 20代   | 男     | 08:00 | エリアA  |
| 30代   | 女     | 08:00 | エリアA  |
| 30代   | 女     | 08:00 | エリアA  |
| 50代   | 男     | 08:00 | エリアB  |
| 50代   | 男     | 08:00 | エリアB  |
| 40代   | 女     | 08:00 | エリアB  |
| ..... | ..... | ..... | ..... |

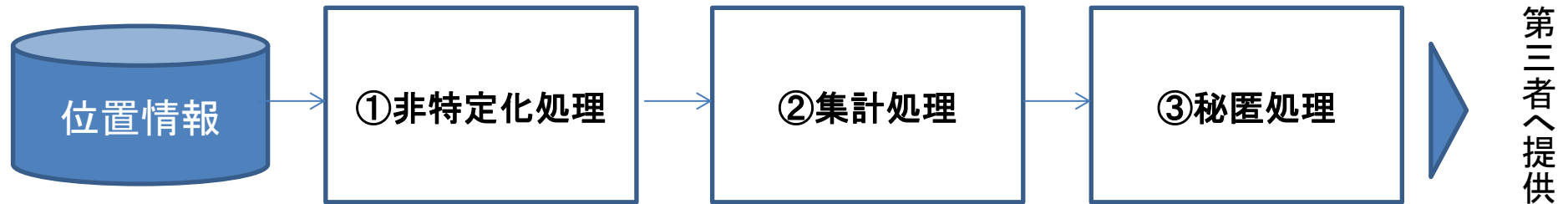


※ 例えば、位置のまとめをしながら全ての属性も加工して、複数人いる状況を作る。IDあり、時点が複数のケースでも拡張可能。

出典: 位置情報プライバシーレポートより

# 有効と考えられる加工を組み合わせた実例： KDDI(観光動態調査)

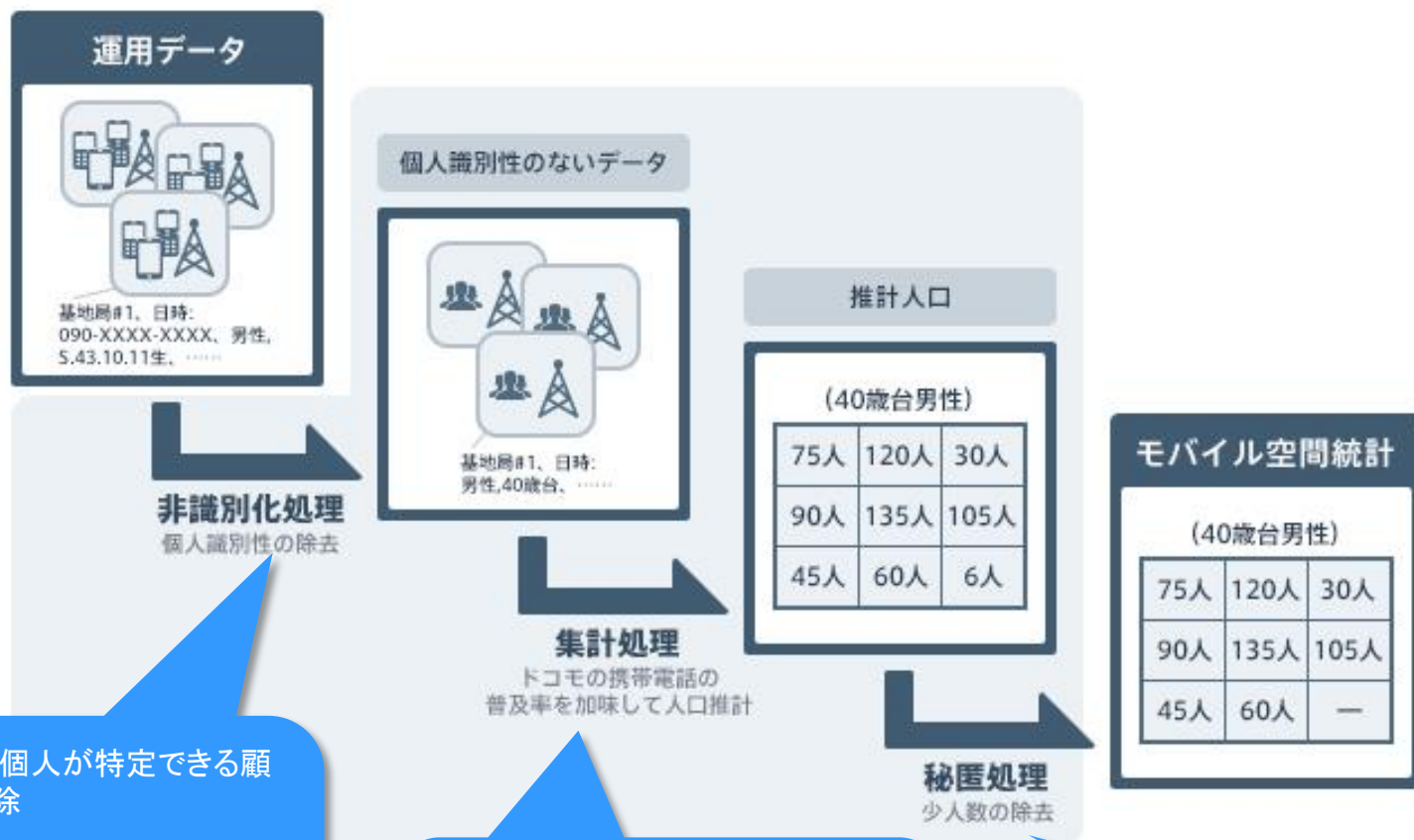
同意の上取得した位置情報の加工プロセスは、大別して3段階の処理に集約される。



|         |   |
|---------|---|
| ①非特定化处理 | 各位置情報ログに対し、IDハッシュ化を行うなど、誰の情報であるかわからない形式に一次加工を施す処理 |
| ②集計処理   | 非特定化处理後のデータを用いた集計処理                               |
| ③秘匿処理   | 集計結果が少数となる場合に表示をマスク化するなどの除外・加工処理                  |



# 有効と考えられる加工を組み合わせた実例：NTTドコモ（モバイル空間統計）



・氏名などの個人が特定できる顧客情報の削除

・電話番号の仮名化

・生年月日から年齢層、住所から居住エリアへの一般化

・基地局位置情報から500mメッシュ等調査ゾーンへの一般化 (図表1に相当)

・時間情報から時間帯への一般化

・少人数の除去による複数化 (図表4に相当)

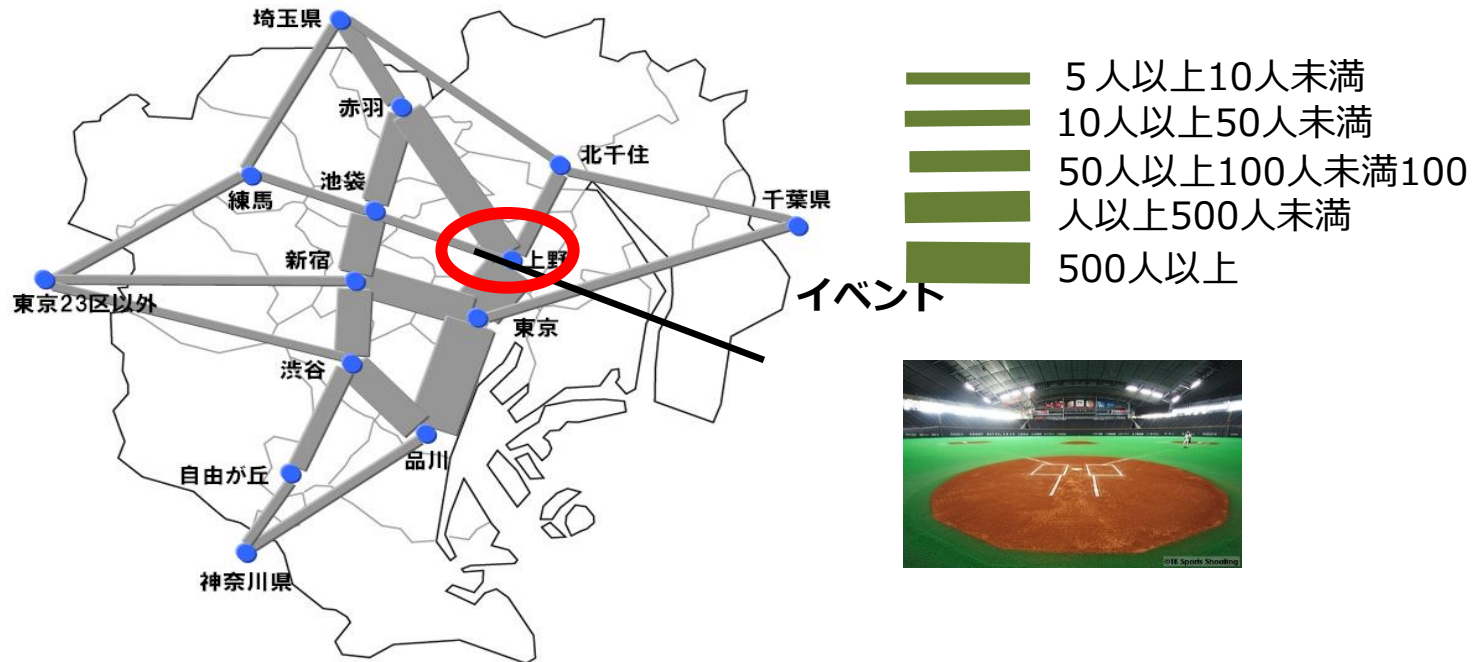
出典：位置情報プライバシーレポートより

# 有効と考えられる加工を組み合わせた結果のアウトプットイメージ

## 位置情報を利用した動線分析の場

分析結果を可視化し、下記のようなイメージを示すことで「十分に匿名化」している安心感を与えられるのではないかと。

例：人口流動調査  
東京近郊のどのエリアからどの程度の人がイベントに参加したか



- ・データは30分単位の取得間隔で収集しています。
- ・各エリアは駅から半径500mの範囲で算出しています。

# 海外の事例(1)

## AT&T Morrystownプロジェクト

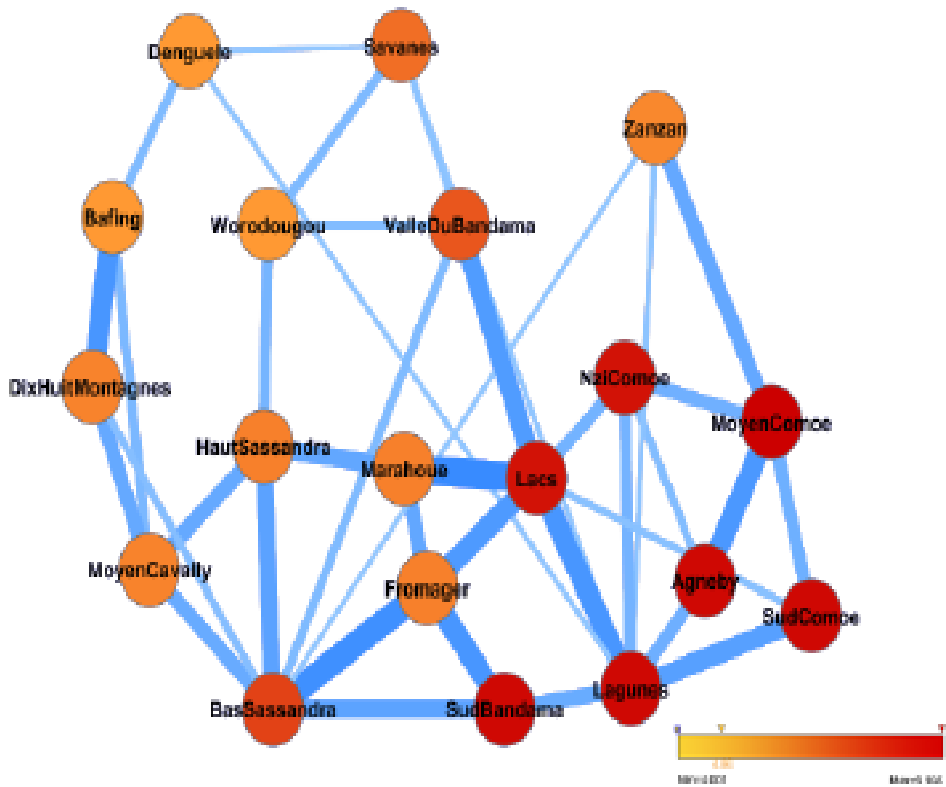
MorrystownはAT&T Researchの近くの小規模な都市  
通勤経路推定、交通状況推定、(通信)トラフィックからの  
イベント検出など多様な分析を展開



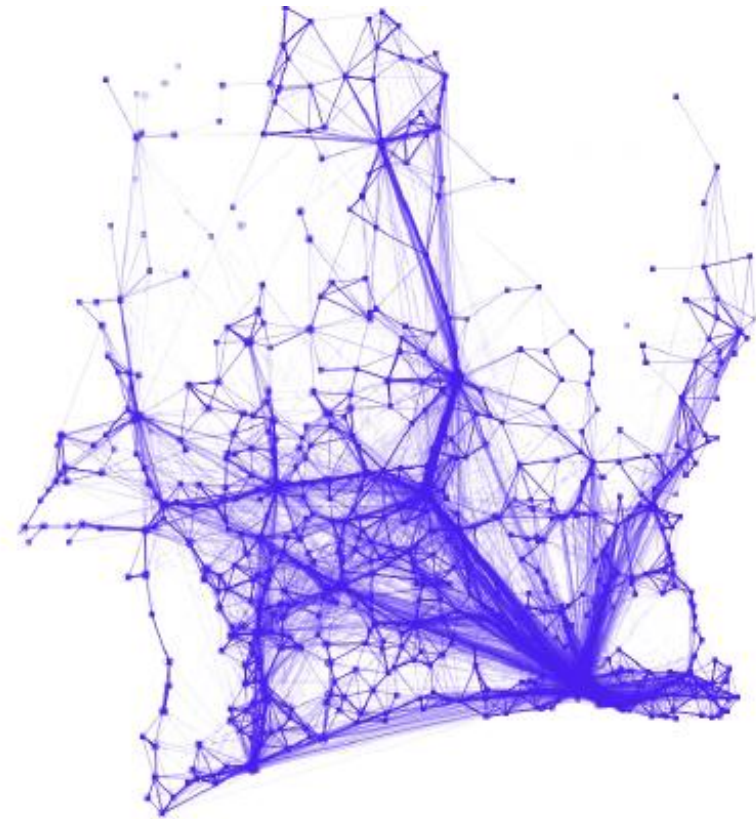
## 海外の事例(2)

### Trg Dositeja Obradovića 6 (セルビアの大学)

- ・ HIVの感染の広がりをモバイルデータを用いてモデル化し検証



SET1: 基地局間のつながりの強さによる伝搬モデル



SET2: 人の動きから感染の経路やポイントからの伝搬モデル