

ビッグデータから見える社会

ー Web/Wi-Fiビッグデータの活用 ー

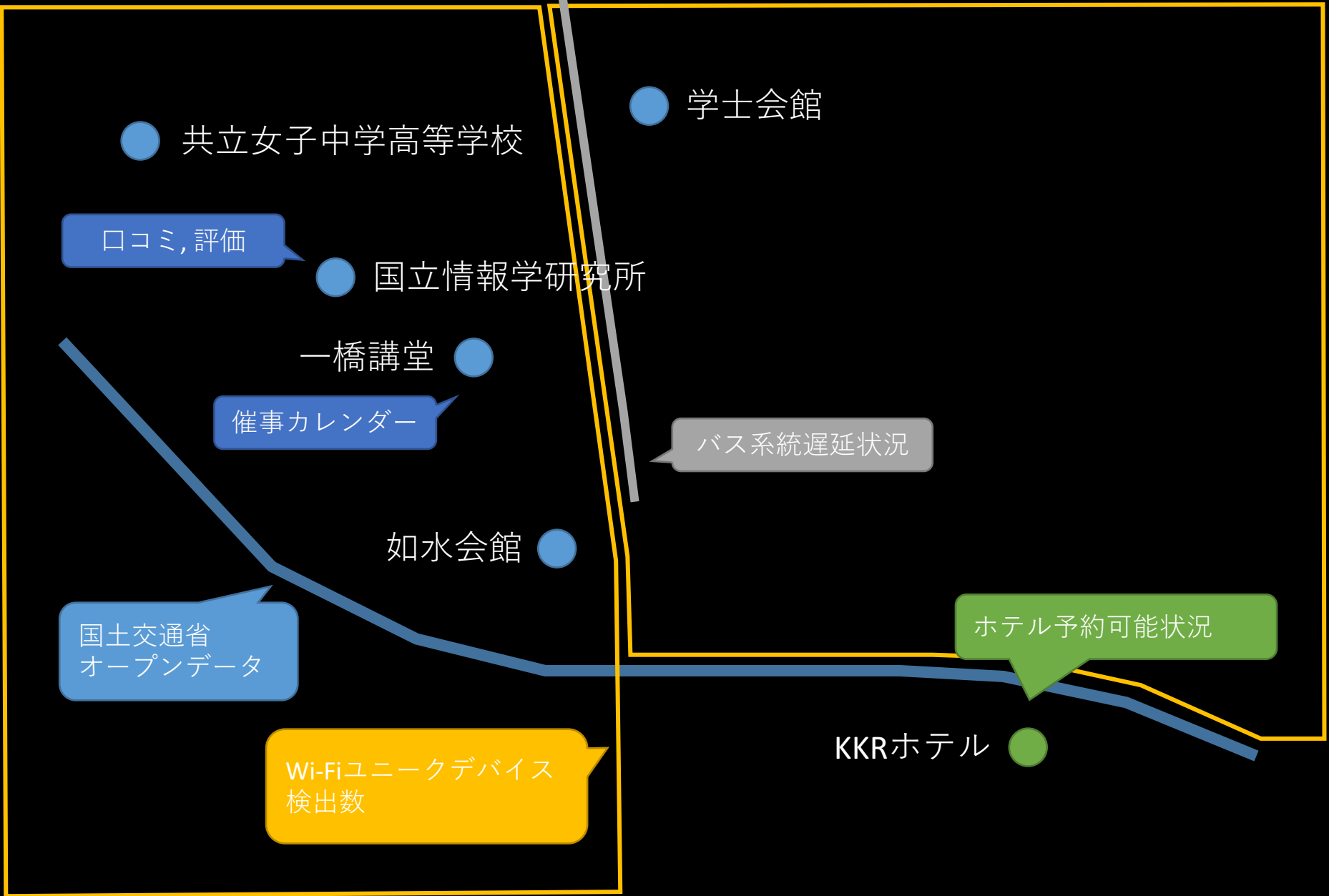
平成29年度 市民講座「情報学最前線」第3回

2017年9月12日@一ツ橋

大学共同利用機関法人 情報・システム研究機構

データサイエンス共同利用基盤施設 社会データ構造化センター 特任研究員

小出哲彰



本日の講演内容

1. Web/Wi-Fiに関するビッグデータ
2. 現実世界の様子を捉え
3. どのように活用するかについて

① データを集める技術

② 分析の方法

③ 最適化への活用

自己紹介

名前

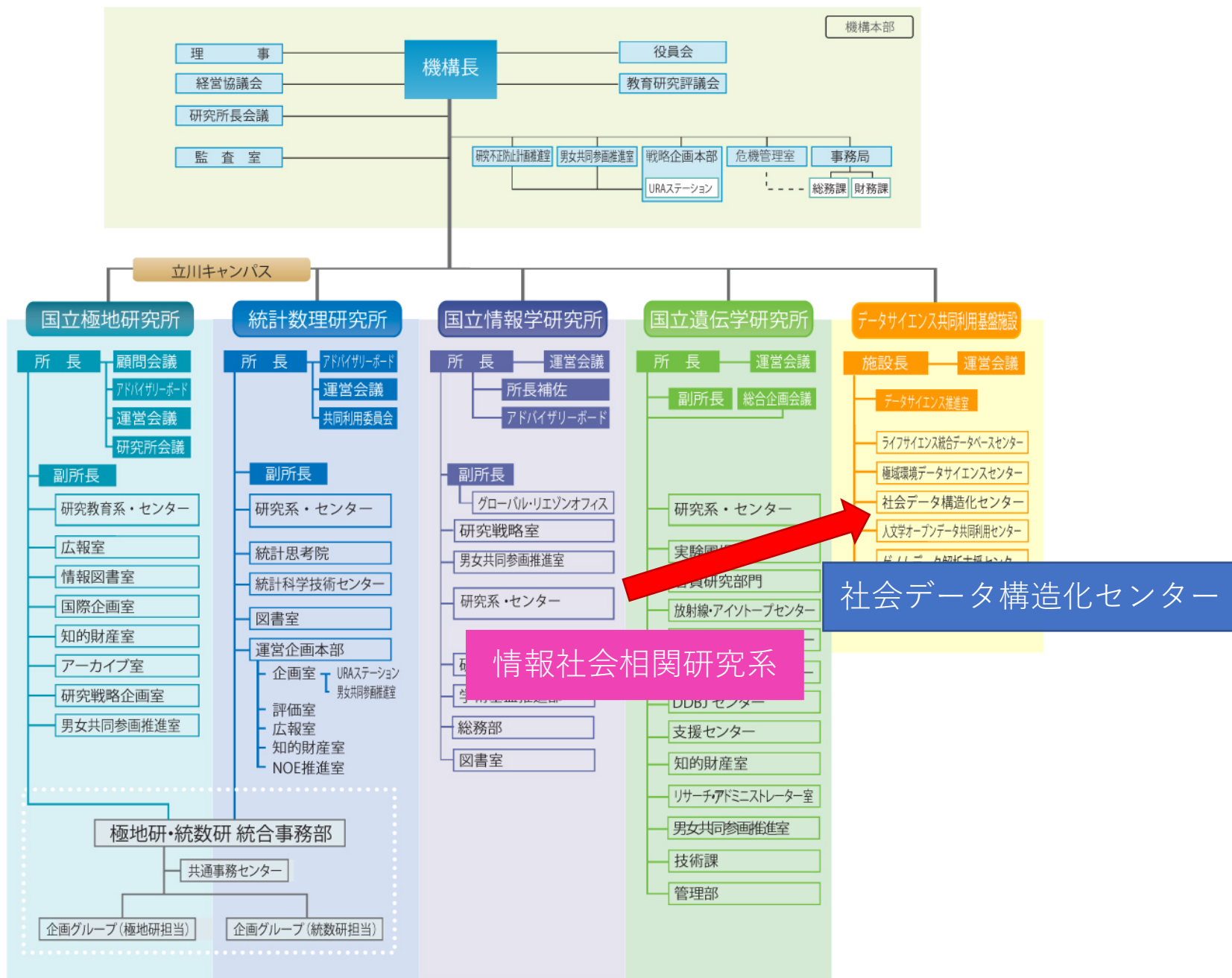
- **小出 哲彰** （こいで のりあき）
- @noriaki_koide （ツイッター）

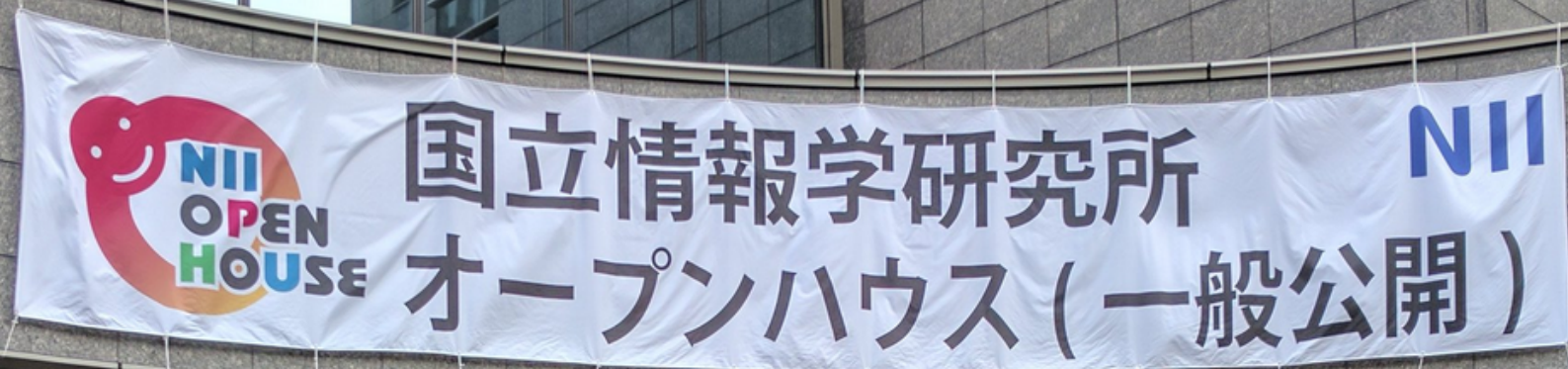
仕事

- **情報・システム研究機構 データサイエンス共同利用基盤施設** 特任研究員
 - 社会データの大学利用の推進
 - ソーシャル・ビッグデータの利活用基盤に関する研究開発に従事
- **株式会社與商会** 代表取締役
 - チャットボットの研究開発
 - 大学への講座提供（公務員・大学院進学）

興味

- 人工知能と教育
- 社会システム





小出哲彰 @noriaki_koide · 6月10日
国立情報学研究所のオープンハウス2017で研究の話をさせていただきました:)
#NIINow

NII プレスリリース

2017/03/14

エビデンスに基づく政策・意思決定を支援／多様なデータを構造化して高速処理／ ソーシャル・ビッグデータ駆動の政策決定支援基盤

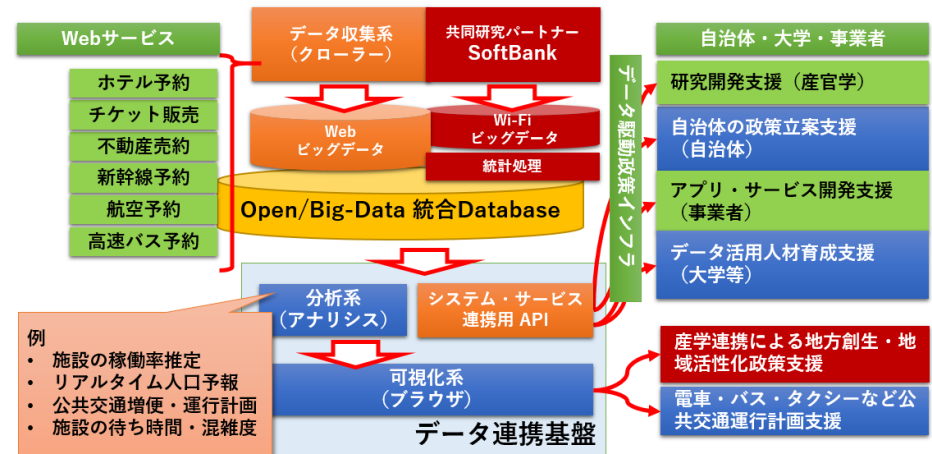
大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 (NII、所長：喜連川 優、東京都千代田区)は3月14日、NII情報社会関連研究系教授・研究主幹 曾根原 登(そねはら・のぼる)の研究チームが開発した「ソーシャル・ビッグデータ駆動の政策決定支援基盤」の詳細について発表しました。

「ソーシャル・ビッグデータ」とは、ビッグデータの中でも特に公共性を有し、社会のさまざまな分野において社会的課題の解決や新たなサービスの創出などに貢献できるものを指します。例えば、インターネットで取得できる宿泊施設や交通機関の予約状況といった社会経済活動データや自治体が収集・提供しているオープンデータなどがこれにあたります。

「ソーシャル・ビッグデータ駆動の政策決定支援基盤」は、こうしたソーシャル・ビッグデータをリアルタイムに連携させることで社会の異なる事象を組み合わせ可視化し、エビデンス(科学的根拠)に基づく合理的な政策決定・意思決定を支援するシステムです。これにより、従来は1週間後、1カ月後に集計された情報を基に行っていた意思決定を現状に即して迅速に実行することが可能となります。

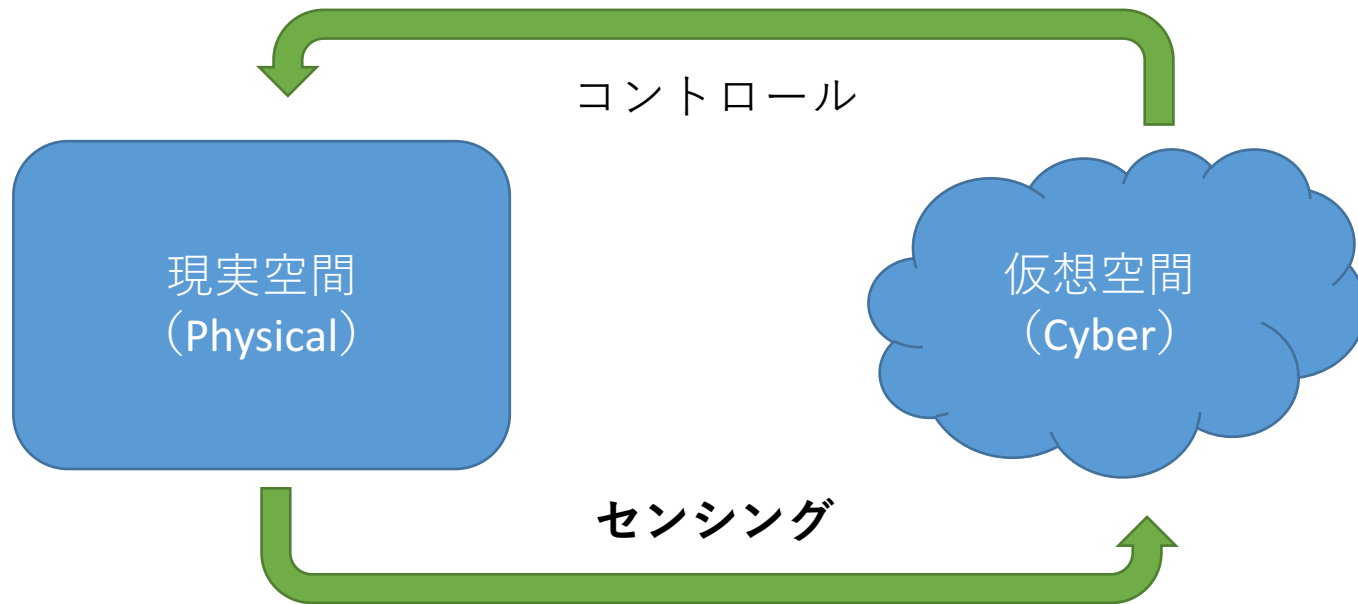
想定される活用例の一つに、地方創生で大きな柱となる観光分野があります。例えば、インバウンド旅行客の動きと宿泊施設の客室提供可能数を組み合わせリアルタイムに把握することで、自治体や事業者は需要と供給のバランスが取れた効果的で即時性の高い施策を実施できるようになります。外国人旅行客が多数集中して来日するオリンピックのような大規模国際イベント開催の際に人の集団の動き(群流)を把握することは、交通機関や雑踏警備の要員配置などの対応策の立案に有効だと考えられます。災害発生時の対応への活用も考えられます。

曾根原教授の研究チームが開発した「ソーシャル・ビッグデータ駆動の政策決定支援基盤」の技術的特徴は、多様なデータの「構造化」と処理速度の「高速化」です。多様なログデータをデバイスごとに整理し、時間的な前後関係に基づいて動線ベクトル・データを生成することで、複雑な群流の解析でも高速に処理できる「TPPP(Trackingid-Timewindow-Place-Pair)手法」を考案しました。



キーワード：Cyber-physical system

サイバー・フィジカル・システム



例

- 送電網のスマートグリッド
- 自動運転車のセルフドライビング・カー (UGV)
- バイタルモニタ (臨床モニタ)
- 組込みシステムの制御 (ロボット)

1. Khaitan et al., "[Design Techniques and Applications of Cyber Physical Systems: A Survey](#)", IEEE Systems Journal, 2014.

センシング (Wi-Fi)
- ヒトの移動を知りたい

免許を必要としない無線局 (= 小電力無線局)

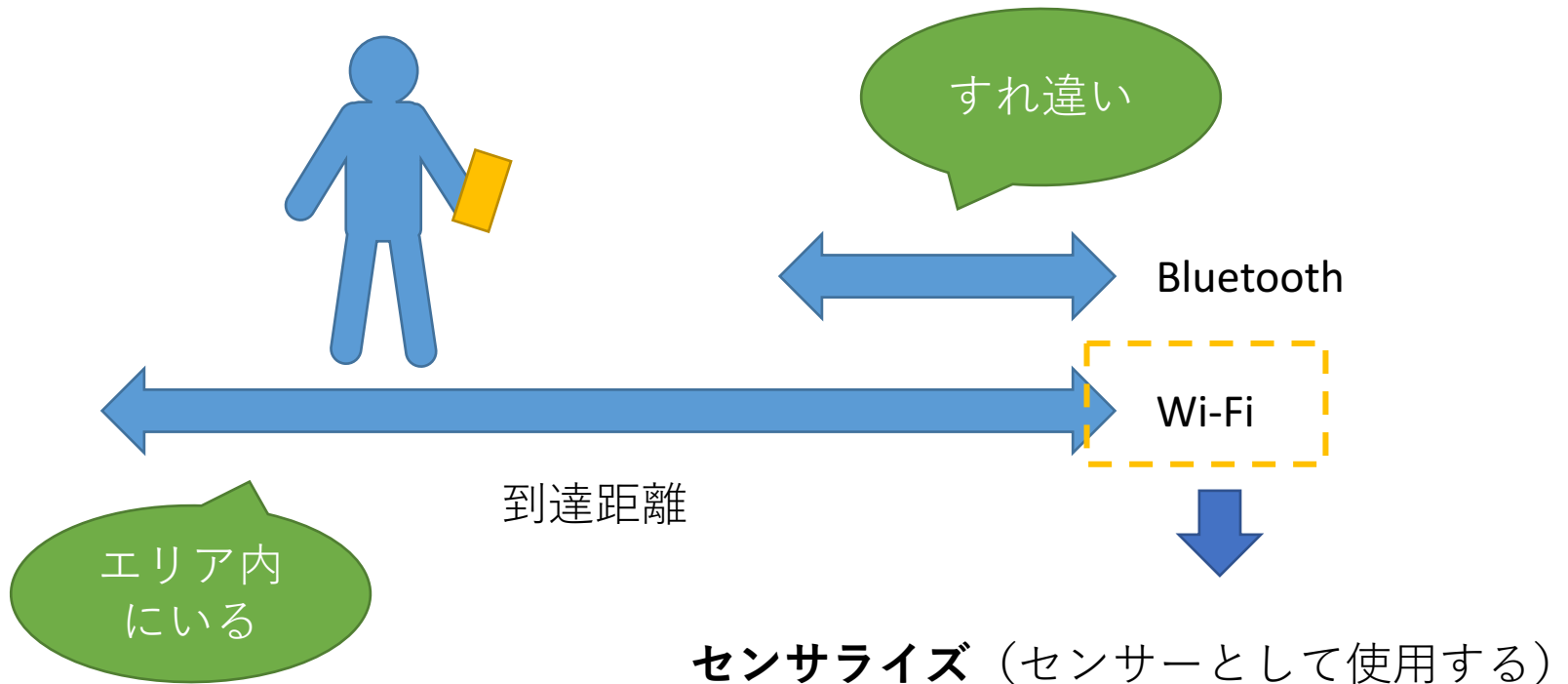
- 出力に上限が定められている

Wi-Fi

- 日本の規格であれば1.0-10.0mW
- 到達距離：障害物のない状態で約100m

Bluetooth

- class2 (スマホ) なら2.5mW
- およそ10m



1. ALSOK: ALSOKの「みまもりタグ」等を活用した“地域の見守り”が国土交通省のモデル事業に選定, https://www.alsok.co.jp/company/news/news_details.htm?cat=2&id2=825
2. Fadel Adib and Dina Katabi.: *See Through Walls with Wi-Fi!*, ACM SIGCOMM'13, Hong Kong, August 2013

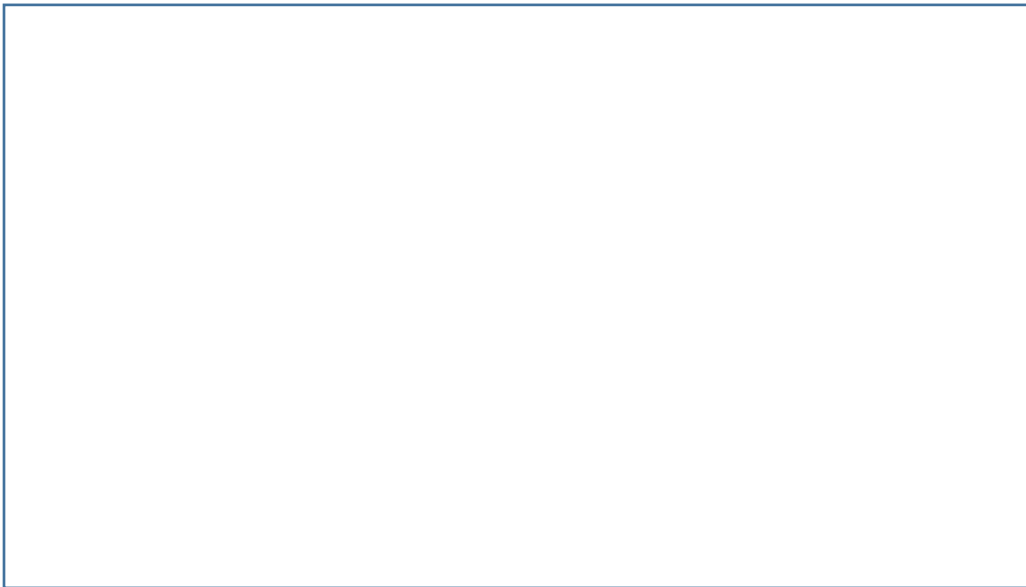
生活に関わる場所にWi-Fiのアクセスポイントが設置されている

- e.g. 図書室, 居酒屋, ホテル, カフェ

キャリア 3 社でも約85万のWi-Fiアクセスポイントが稼働

- 設置箇所や設置数に関する情報はホームページから確認できる

ちょこっとメモ

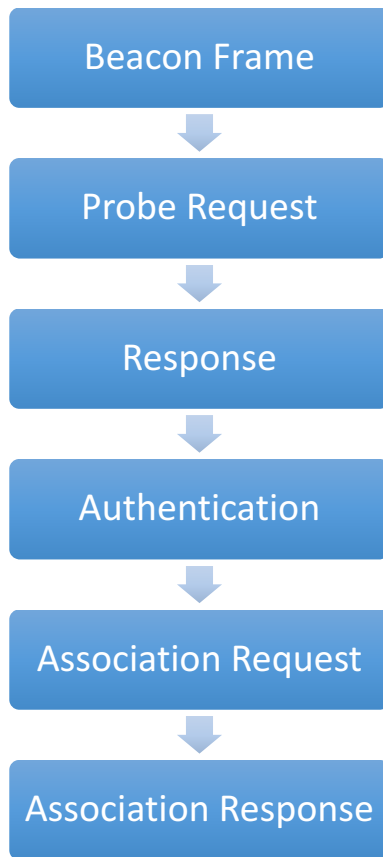


Keyword: 現在位置, GPS, Wi-Fi



Google street view car by Mike Cattell

センシングの前に Wi-Fiの仕組み



Wi-Fiが繋がるまでの
アクセスポイントと端末のやり取り

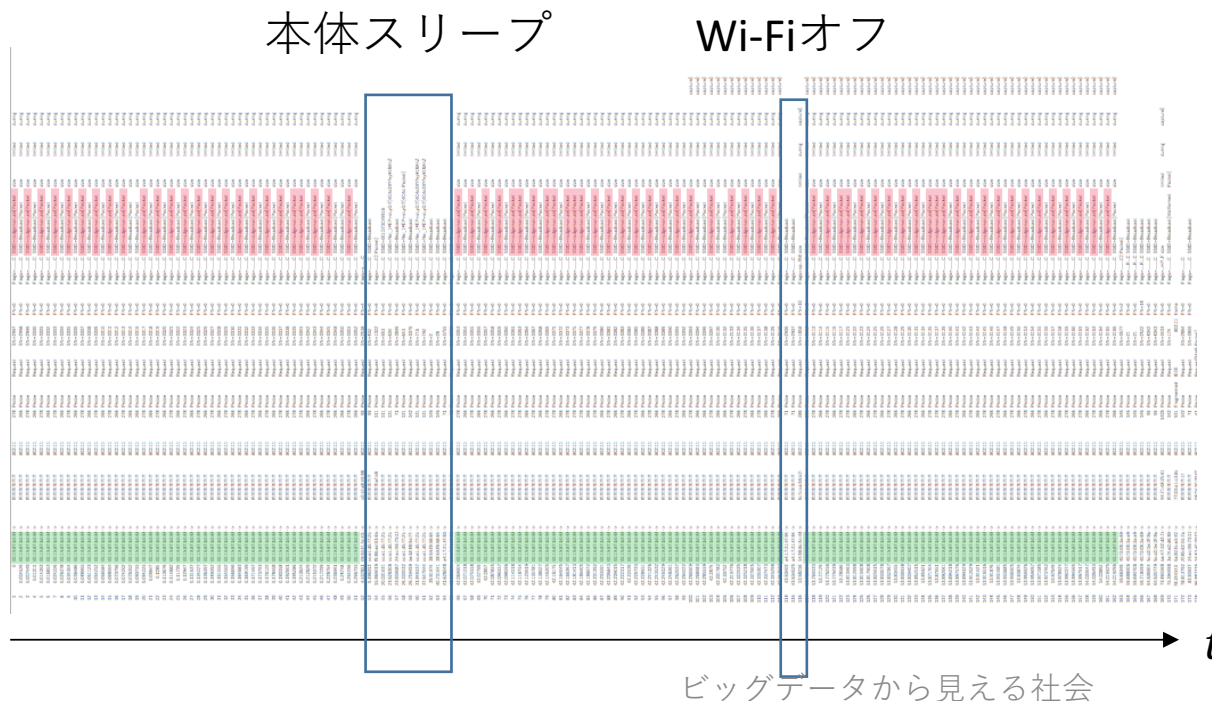
- Wi-Fiが繋がるまでに、アクセスポイントと端末の間でパケットのやり取りが行われる。
- これを記録することで、端末のセンシングを行う。

パケット名

- **wireshark**などのネットワークプロトコルアナライザで、端末とアクセスポイントの間でどのようなパケットのやり取りが行われているかを眺める。
- 自分の端末を検出してみる。

テスト – Android 4.2.2

- Wi-Fi信号(Probe Request)を使用して、特定のAndroidデバイスをMACアドレスで追跡することを試みた。
- スリープやWi-FiのオンオフによりMACアドレスは変化せず、定点観測による追跡が可能。

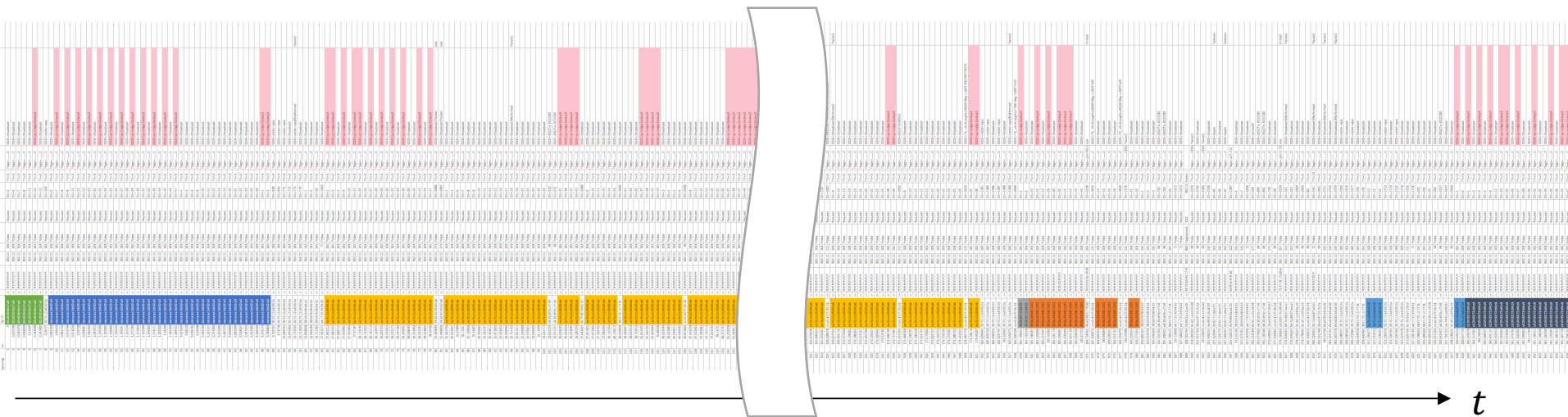


テスト – iOS 9.2

- Wi-Fi信号(Probe Request)を使用して、特定のiOSデバイスをMACアドレスで追跡することを試みた。
- 特定のiOSデバイス（ピンク色で網掛け）をSSID一覧に持つ小出のデバイス。Probe Requestを送信している。
- スリープやWi-FiのオンオフによりMACアドレスが変化した。

緑 → 青 → 黄 → 灰色 → 橙色 → シアン → ネイビー

\$ tshark -lni en1 -s 256 type mgt subtype probe-req



Wi-Fiアクセスポイントを用いた 端末のセンシング（トラッキング）

Association
00:00:00 AAAAAAA
00:00:01 BBBBBBB
00:00:02 CCCCCC
00:00:03 DDDDDDD
...

アクセスポイント(1)

Association
00:00:00 EEEEEEE
00:00:01 FFFFFFFF
00:00:02 AAAAAAA
00:00:03 GGGGGGG
...

アクセスポイント(2)

Association
00:00:00 HHHHHHH
00:00:01 KKKKKKKK
00:00:02 BBBBBBB
00:00:03 CCCCCC
...

アクセスポイント(3)

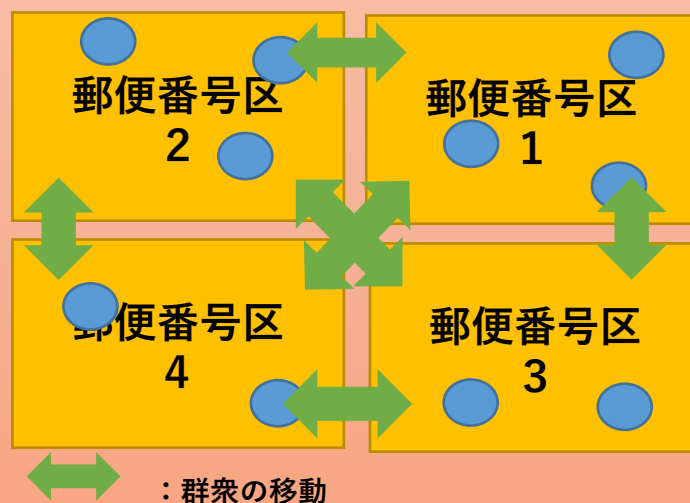
- 各アクセスポイントにおいて**Association**された時刻と端末の**MAC**アドレスを記録
- それぞれのログデータを結合することで、特定の端末をトラッキングすることができる

Wi-Fi アクセスポイントのセンサー化



どこから来てどこへ行ったか、 群単位の移動の可視化

- 事前同意済みインバウンド観光客の統計化 (Statistical processing of inbound tourists with prior consent)



センシング (Web)

- 施設の状態などを知りたい

開催プログラム

*各回ごとに受講できます。

第1回 2017年7月12日(水) 18:30-19:45 <文字通訳あり>



YouTube

資料

レポート

やる気を引き出す人工知能

—個性を活かして学習意欲を高めるAI—

坂本 一憲

【概要】

近年、オンライン講座サービスなど様々な教育ソフトウェアが普及しつつあり、それに伴い、学習効果を高めるために学習意欲を引き出す方法が開発されています。多くのサービスは全ての学習者に画一的な方法を提供していますが、学習者は様々な個性を持っており、個性によって有効な手法が異なることが、心理学研究で明らかになっています。

私たちは、人工知能技術と学習者の心理的な傾向を分析する手法を組み合わせることで、学習者ごとに最適な方法を提供する技術と本技術を搭載したソフトウェアを開発しています。本講座では、本技術の開発を通じて明らかになった知見を共有するとともに、人間の意欲について皆で考えるきっかけを作ることを目指しています。

第2回 2017年 8月25日(金) 18:30-19:45 <講義映像の公開はありません>



レポート

Q&A

ダイヤモンドと量子情報

—テレポーテーションから量子認証まで—

小坂 英男(横浜国立大学 大学院工学研究院 教授)

国立情報学研究所 量子情報国際研究センター

新学術領域「ハイブリッド量子科学」研究メンバー

【概要】

ダイヤモンドは最も貴重な宝石として知られますが、実は量子と呼ぶとても壊れやすい状態を長く保持する量子結晶としても貴重です。ダイヤモンドの欠陥に窒素を封じ込めた窒素空孔(NV)中心は、光の最小単位である光子の量子状態を電気の最小単位である電子へと量子テレポーテーションの原理で転写し、磁気の最小単位であるスピンとして保存する天然の量子メモリーとなります。このようなダイヤモンドの特殊能力は、モノのインターネット(IoT)の安全性を量子力学

```

399 <h4>開催プログラム</h4>
400 <p>*各回ごとに受講できます。</p>
401 <table class="spTate"><colgroup> <col width="25%" /> <col /> </colgroup>
402 <tbody>
403 <tr id="1st"><th colspan="2">第1回 2017年7月12日(水) 18:30-19:45 <文字通訳あり>
404 <tr>
405 <td>
406 <p>
407 </td>
408 <td><strong>やる気を引き出す人工知能</strong><br />—個性を活かして学習意欲を高めるAI—
409 <div><strong>【概要】</strong></div>
410 <div>近年、オンライン講座サービスなど様々な教育ソフトウェアが普及しつつあり、それに伴い、学習効果を高めるために学習意欲を引き出す方法が開発されています。多くのサービスは全ての学習者に画一的な方法を提供していますが、学習者は様々な個性を持っており、個性によって有効な手法が異なることが、心理学研究で明らかになっています。
411 </div>
412 </td>
413 <tr id="2nd"><th colspan="2">第2回 2017年 8月25日(金) 18:30-19:45 <講義映像の公開はありません>
414 <tr>
415 <td>
416 <td><strong>ダイヤモンドと量子情報</strong><br />—テレポーテーションから量子認証まで—
417 <div><strong>【概要】</strong></div>
418 <div>ダイヤモンドは最も貴重な宝石として知られますが、実は量子と呼ぶとても壊れやすい状態を長く保持する量子結晶としても貴重です。ダイヤモンドの欠陥に窒素を封じ込めた窒素空孔(NV)中心は、光の最小単位である光子の量子状態を電気の最小単位である電子へと量子テレポーテーションの原理で転写し、磁気の最小単位であるスピンとして保存する天然の量子メモリーとなります。このようなダイヤモンドの特殊能力は、モノのインターネット(IoT)の安全性を量子力学
419 </div>
420 </td>
421 <tr id="3rd"><th colspan="2">第3回 2017年9月12日(火) 18:30-19:45</th></tr>
422 <tr>
423 <td>
424 <td><strong>ビッグデータから見える社会</strong><br />—Web/Wi-Fiビッグデータの活用からセキュリティまで—
425 <div><strong>【概要】</strong></div>
426 <div>Web、Wi-Fiは私たちの生活でより身近なものになりました。インターネット上に
427 </div>
428 </td>

```

スクレイピング

= Webページから必要な情報を抜き出す

1. 加藤耕太: Python クローリング&スクレイピング データ収集・解析のための実践開発ガイド, 2017

Webページ（宿泊施設予約）

日	月	火	水	木	金	土
7/30 -	7/31 -	1 -	2 -	3 ×	4 ×	5 ×
6 4室 10,834	7 ×	8 ×	9 ×	10 2室 12,038	11 ×	12 ×
13 5室 11,389	14 5室 11,389	15 5室 11,389	16 5室 11,389	17 5室 11,389	18 3室 11,389	19 ×
20 5室 10,834	21 3室 11,389	22 ×	23 ×	24 ×	25 3室 12,686	26 ×
27 5室 10,834	28 1室 11,839	29 1室 11,839	30 3室 11,839	31 1室 11,839	9/1 ×	9/2 ×

スクレイピング



Webサイトの掲載情報
（カレンダー）をもとに
メタデータを作成する

メタデータの作成においては、宿泊プランの名称がそのまま記録される箇所があれば、次のようなルールにもとづいて記録が行われる

- 「4室」なら「4」
- 「×」なら「0」
- 「○」なら「>=10」

メタデータ

```
--- start ---
1435,20131116,2,1628519,wa10,その日の仕入れ素材にこだわった
★料理長自慢京会席★
くつろぎの和室10畳(バス・トイレ(ウォシュレット)付)
11/16,0
11/17,>=3,25200
11/18,>=3,25200
11/19,>=3,25200
11/20,>=3,25200
11/21,>=3,25200
11/22,0
11/23,0
11/24,>=3,25200
11/25,>=3,25200
11/26,0
11/27,0
11/28,>=3,25200
11/29,>=3,25200
11/30,0
12/01,>=3,25200
12/02,>=3,25200
12/03,0
```

キーワード：正規表現, Xpath, CSSセレクター, 深層学習

Web

クローリング = Webからデータを収集してくること

キーワード：クローラー, スパイダー, ボット

Webクローラーを用いた 宿泊施設予約状況のセンシング

- クローラーで取得してきたWebページのデータを**スクレイピング**することで、必要な情報を抜き出す。

日	月	火	水	木	金	土
7/30 -	7/31 -	1 -	2 -	3 ×	4 ×	5 ×
6 4室 10,834	7 ×	8 ×	9 ×	10 2室 12,038	11 ×	12 ×
13 5室 11,389	14 5室 11,389	15 5室 11,389	16 5室 11,389	17 5室 11,389	18 3室 11,389	19 ×
20 5室 10,834	21 3室 11,389	22 ×	23 ×	24 ×	25 3室 12,686	26 ×
27 5室 10,834	28 1室 11,839	29 1室 11,839	30 3室 11,839	31 1室 11,839	9/1 ×	9/2 ×

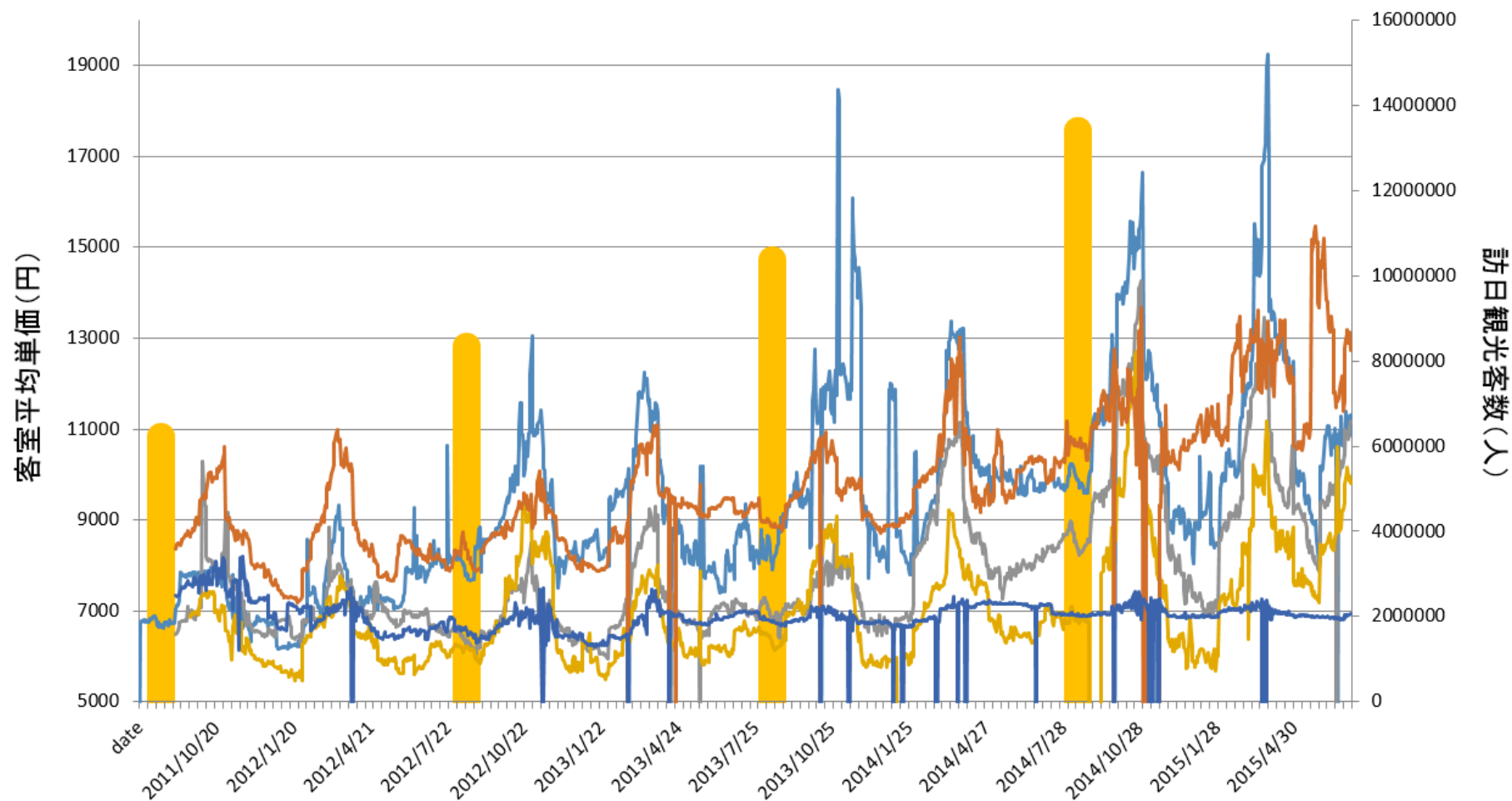


全く予約ができない日付
すなわち客室稼働率が100%



客室数だけ滞在されている
様子が分かる！！

Webクローリングにより収集されるデータ



訪日観光客数の推移と客室平均単価

データを解析する (Wi-Fi)

Estimation of National Tourism Statistics Based on Wi-Fi Association Log Data

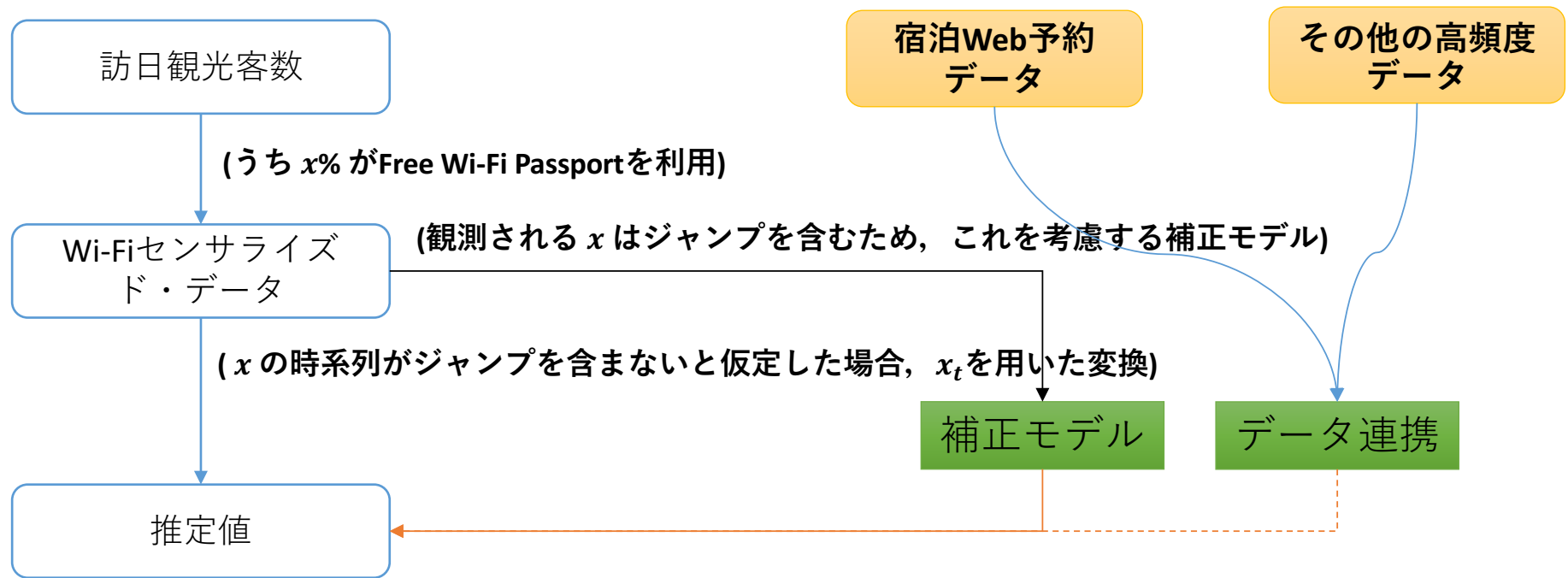


図. Wi-Fiセンサライズドデータを用いた訪日観光客数の推定

- 一部の訪日観光客が国内でWi-Fiを利用し、ログ上の接続デバイスとして観測されることを利用して、観光統計のリアルタイム化を行う。

データ

研究に利用されるデータ（JNTO）

- 日本政府観光局（以降，JNTO）による訪日外国人に関する統計データ
- 「JNTO のホームページ > 統計情報」より一般に入手可能
- 各月について，どの国から何名の訪日があったかを記録
- 観光統計に関する多様なデータあり

データ

今回の検討では，2015 年 7 月から 12 月について

- Wi-Fi ログデータからハッシュ化された MAC アドレスと属性情報（国籍）にもとづいて，各月での各国に対するユニークなデバイス数を抽出
- JNTO 統計データから各月での各国からの訪日外国人数
- それぞれの月次データについて，月次間での検出数および人数カウントデータの伸び率（成長率）を得る

使用する国籍の選定

国籍

Wi-Fi ログデータの提供元となるサービスの利用規約を記述する言語（英語，中国語，韓国語）を公用語とする国を選択した

- 韓国，中国，台湾，香港，シンガポール，
- フィリピン，英国，米国，カナダ，
- 豪州，ニュージーランド

（ただし米国およびカナダについては，Wi-Fi ログデータの属性情報付与方法の特性から一まとめとする）

利用規約

[英語の利用規約をみる](#) 

[中国語（簡体）の利用規約をみる](#) 

[中国語（繁体）の利用規約をみる](#) 

[韓国語の利用規約をみる](#) 

仮説

モデル

Wi-Fi ログから得られる検出数値と JNTO 統計データによる訪日外国人観光客数の間に次の関係を仮定する.

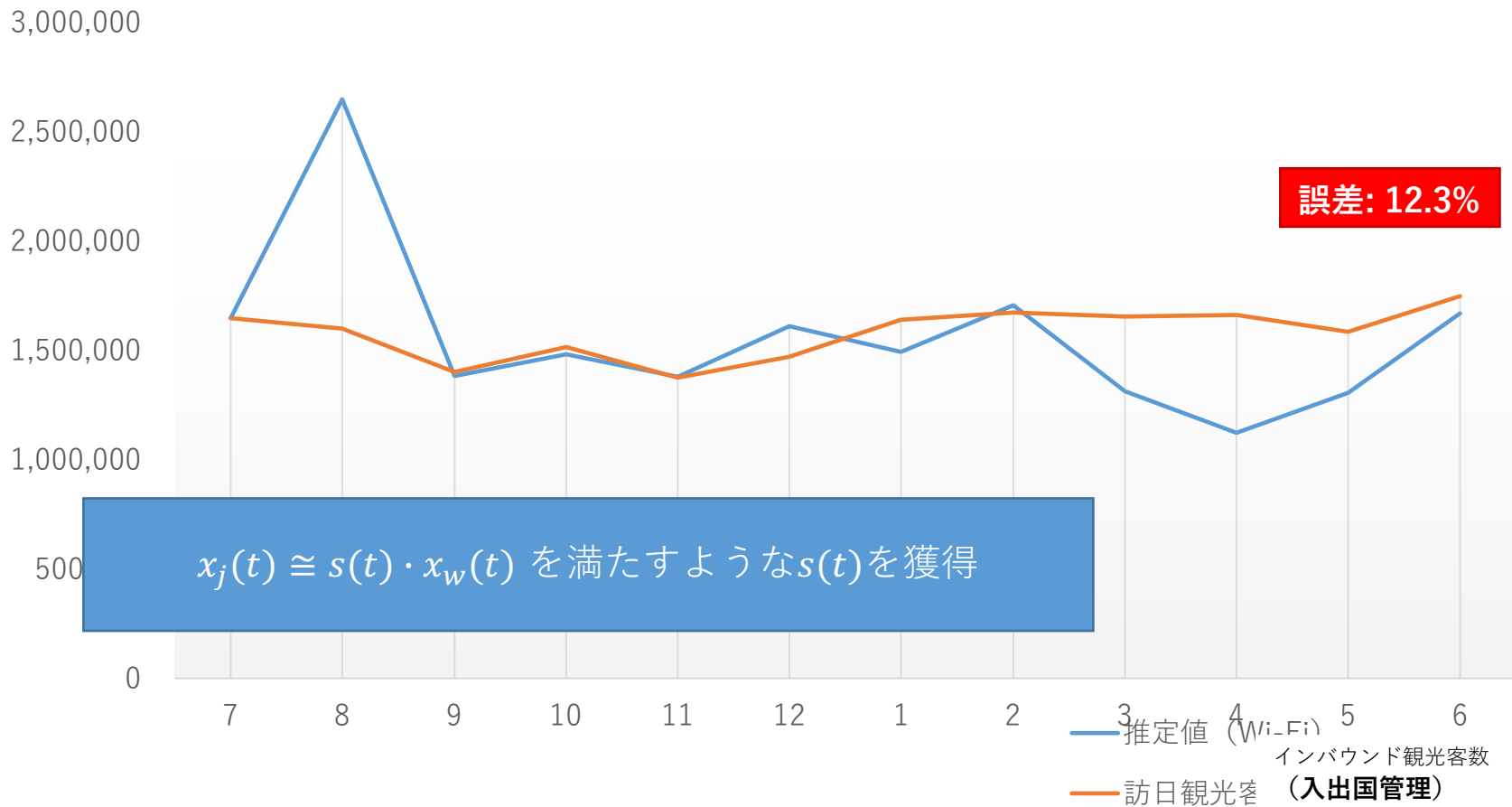
$$W_t = s_t X_t \quad (2)$$

ただし

- W_t を t 期における Wi-Fi ログのユニークデバイス数の検出数
- X_t を t 期における訪日外国人観光客数の実測値 (JNTO 統計などにより遅れて報じられる値)
- s_t を t 期における訪日外国人観光客に占める同サービス Wi-Fi 利用者の割合 (s_t はサービス提供者の利害関係により変動するような確率変数で観測できない)

伸び率については

$$\frac{W_t}{W_{t-1}} = \frac{s_t X_t}{s_{t-1} X_{t-1}} \quad (3)$$

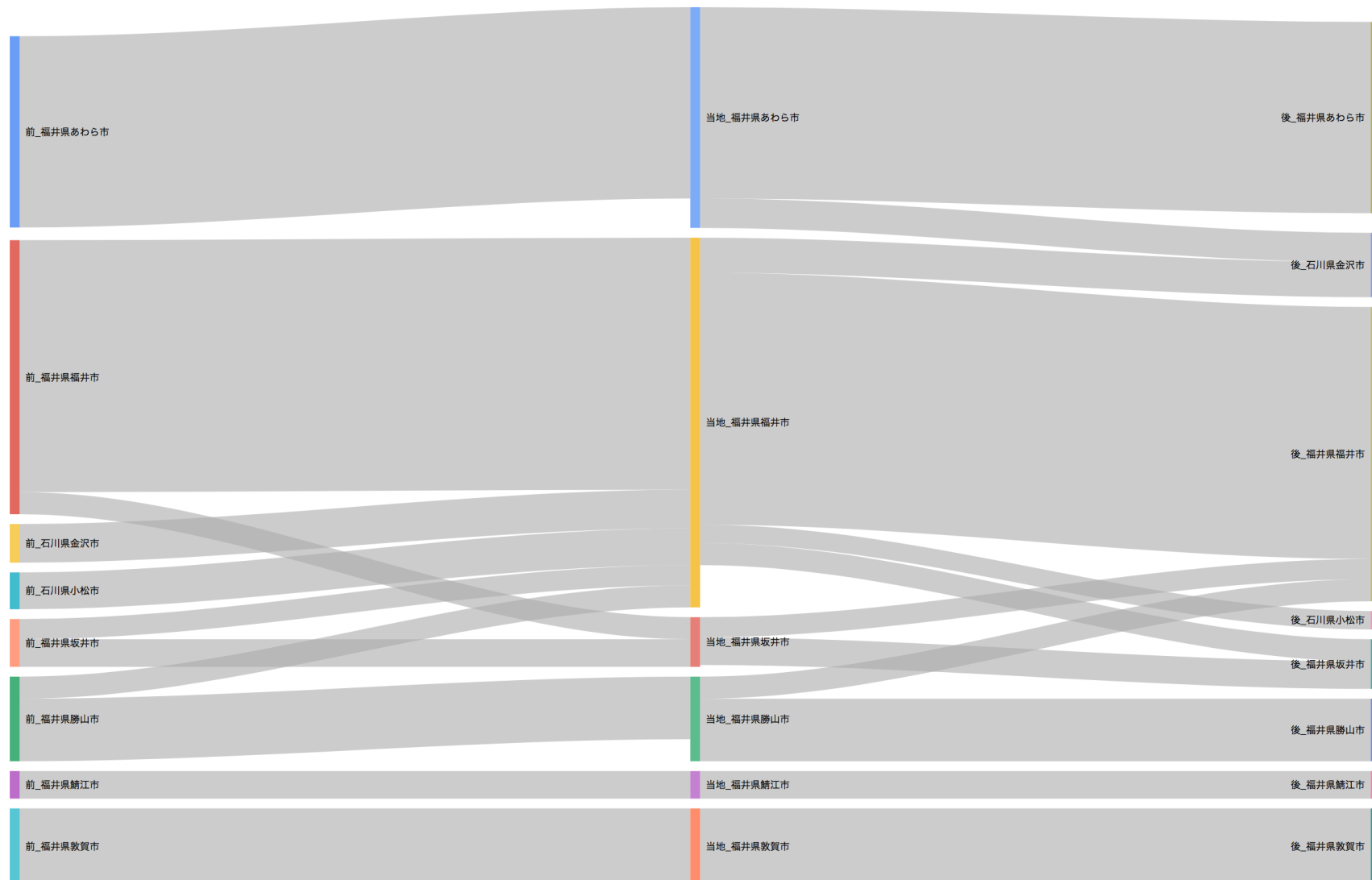


※ Wi-Fiビッグデータ

日次、時間ごとの推定も可能 > 統計調査のリアルタイム化

1. 和合 肇, 萩原淳一郎: **R**によるベイジアン動的線形モデル, 統計ライブラリー, 2013.
2. **Noriaki Koide**, Yu Ichifuji, Hideki Yoshii and Noboru Sonehara, “*Estimation of National Tourism Statistics Based on Wi-Fi Association Log Data*,” 2016 IEEE International Conference on Big Data, Washington D.C., USA, 5-8 Dec. 2016.

福井県内各市ごとの流入 サンキー (2015年7月～2016年6月) 前->当地->後 :福井県



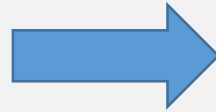
双曲線関数を用いた補正モデル

[新宿1丁目:夜]

$$c(t, s) = S(t) \tanh(\alpha w w_{max}^{-1})$$

〒160-0022 @19:00:00

ユニークデバイス検出数：**23**



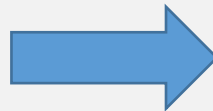
〒160-0022 @19:00:00

推定人数：**1636**

[博多駅南:昼]

〒812-0016 @12:00:00

ユニークデバイス検出数：**1**



〒812-0016 @12:00:00

推定人数：**5.78**

- 1ヶ月間を平滑化した値

データを解析する (Web)

- 宿泊施設予約状況の例

宿泊施設の稼働率の推定

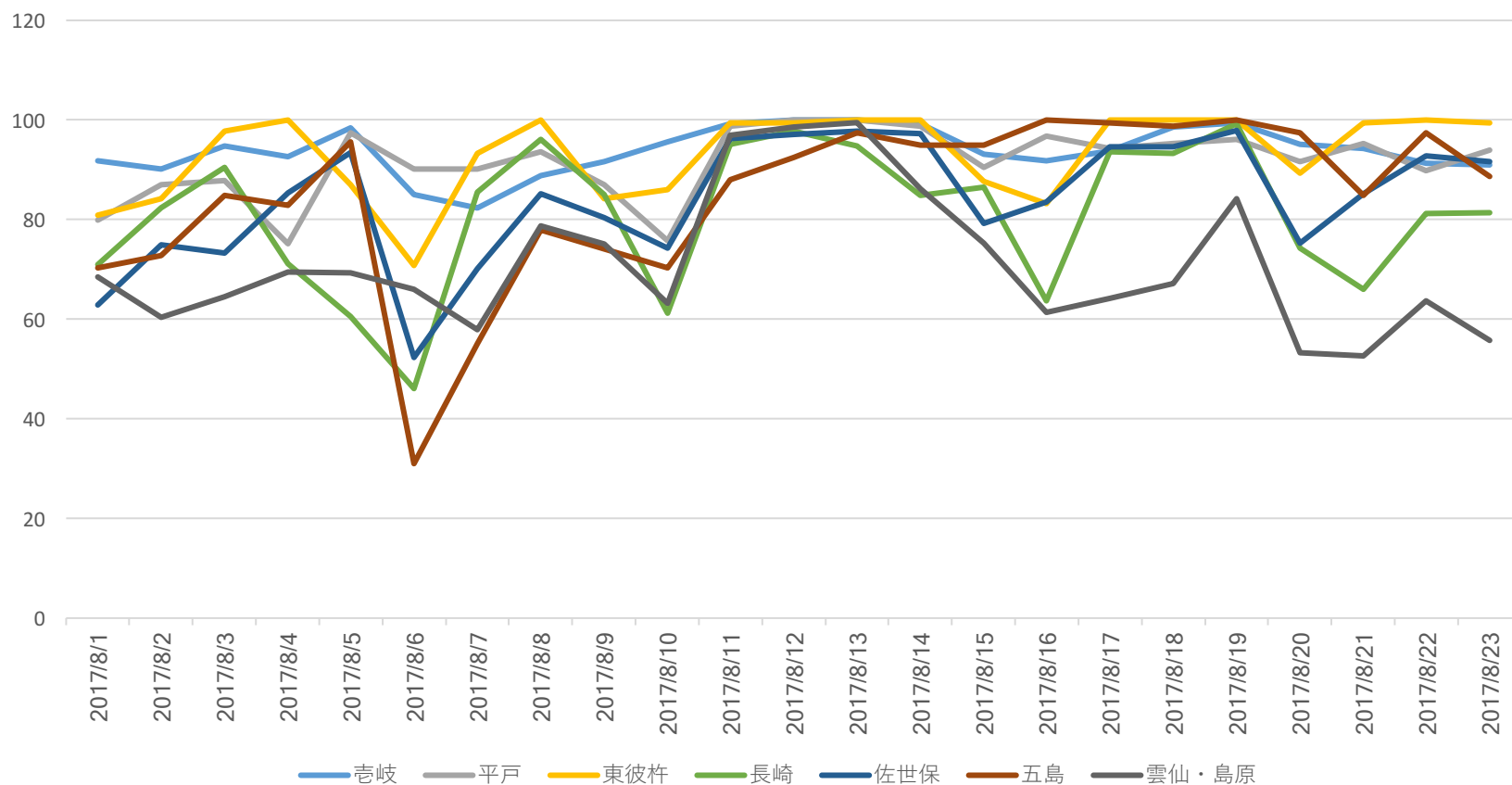
- 自治体・行政が調査を行なっている「宿泊旅行統計調査」を近似したい。
 - 各ホテルに**稼働状況**をヒアリング
 - 調査コストが高い
- Webから取得できるデータをもとに**ホテルの稼働状況を推定**する。
 - 7%の誤差で推定（月次）

1. 情報・システム研究機構：「Webデータからホテルの料金と空室数を予測する！ 『ビッグデータ駆動の観光・防災政策決定支援システムの研究開発』」（2013年3月6日）

アルゴリズム

1. ホテル・対象日について、データの差分を計算する
2. 部屋IDが同じものを選択し、差分0～10を1つのデータとして、他のデータとの比較を行う。
 - A) 同一のものがあれば、当日の予約可能プラン数が大きいほうを残す
 - B) 同一のものがなくなるまで、(A) を繰り返す
 - C) すべての部屋IDに対して、(A) (B) を繰り返す
 - D) 残ったものの予約可能プラン数を空室数として出力する。
 1. 空室数がホテルの全部屋数よりも大きい場合、(A) の比較でj箇所以下が異なっても同一とみなす、この数値jを設定する場所を追加する。
3. ホテルテーブルから、部屋数を抽出する。
4. 2, 3 の部屋数と空室数から対象日の稼働率を計算する

地域の宿泊施設の稼働率



例：宿泊施設の稼働率とヒトの移動

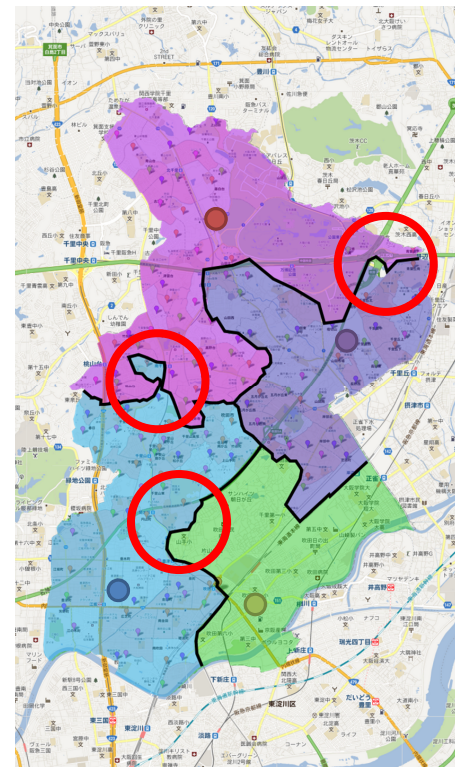
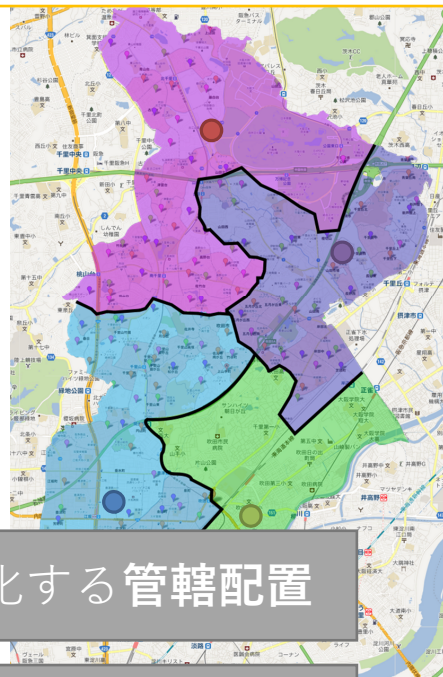


データを活用する

活用1: 消防署管轄についての最適化

現実問題 「緊急時に機敏に対応できる消防署の管轄は？」

- 各住所には地域の消防署が**割当て**られている
- 供給側である消防署は限定されている
- 実際に通行する経路を考慮する必要がある



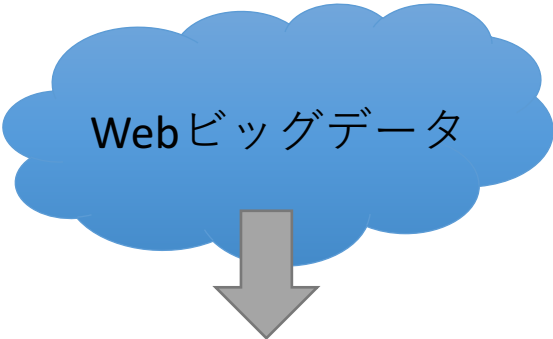
厚生損失を最小化する管轄配置

地の利を数値化（固有ベクトル）

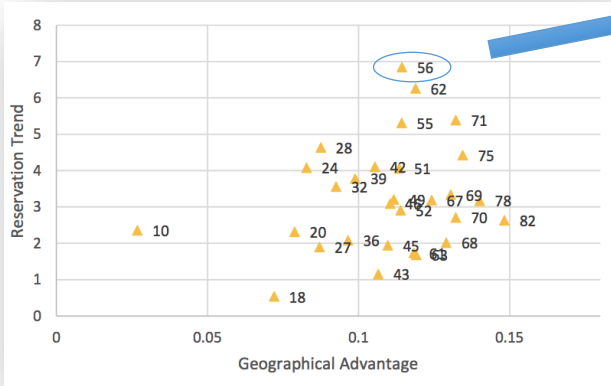
市内の住所の関係をグラフ化

国立情報学研究所

活用2: 宿泊施設の予約傾向と地理特性, 効率性評価



	DMU1	DMU2	DMU3	DMU4	...	DMU56	...	DMU85
Number of Rooms	37	100	161	82	...	46	...	482
Price of Staying	4500	3410	3300	4200	...	1300	...	3500
Geographical Advantage	0.0127	0.0141	0.0200	0.0231	...	0.15596	...	0.1026
Review Point	4.09	4.17	3.83	3.62	...	4.21	...	4.2
Reservation Trend	2.429	2.8505	3.9840	1.9439	...	6.8434	...	4.5245
D-Efficiency Value	0.8071	0.9209	0.8345	0.5921	...	0.4102	...	0.5031



DEA（包絡分析）による
効率性評価



図2. 国道（緑）沿いのホテルがより優
位な地利値を有する

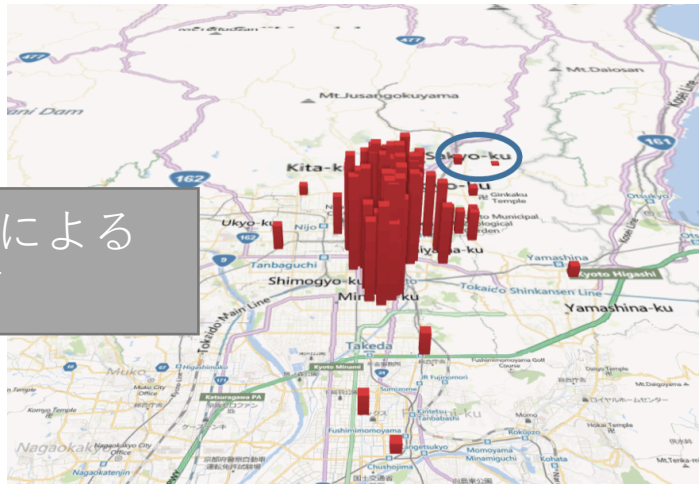
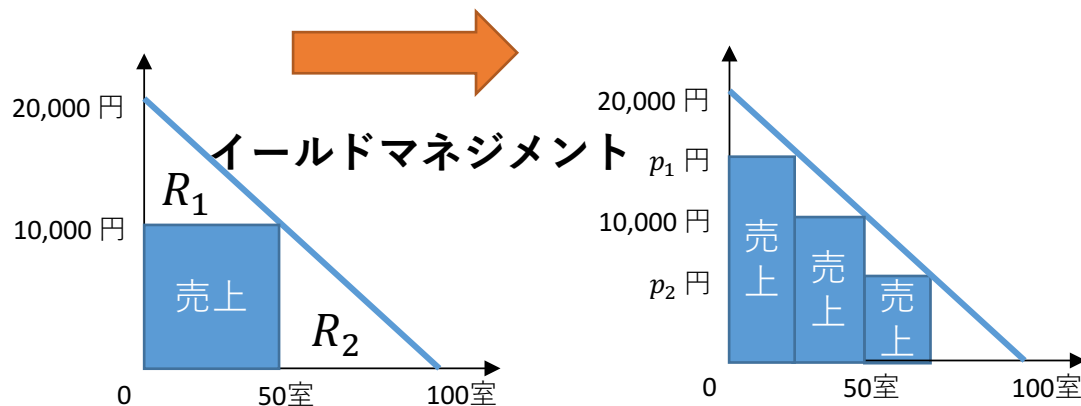


図1. 京都に点在するホテルが持つ地利
値（移動時間により算出）

活用: 客室単価の価格設定に関する意思決定



R_1 : 価格を高めにも設定しても売れた部屋 p_1 : 価格を気にしない宿泊者

R_2 : 価格を低く設定していれば売れた部屋 p_2 : 価格に敏感な観光客

※ 青線 (—) は需要曲線を表す

現実問題 「1室しかもたないホテルは客室単価をどう設定すれば良いか？」

- 既存のイールドマネジメントに対して、
- Webビッグデータを活用したアプローチ

提案手法

1部屋しかない宿泊施設

1. 価格設定を行いたい宿泊施設に類似した複数の宿泊施設を1つの宿泊施設と看做し、Web上でプランの予約傾向を観察する。
2. この傾向情報に基づいて、期待収入に関する動的最適化問題を解く。

$$V_t(x) = \sum_{i=1}^m p_{i,t} \max\{r_i + V_{t-1}(x+1), V_{t-1}(x)\} + p_{0,t} V_{t-1}(x)$$

地域・ホテルのイールドマネジメント

従業員**10**名以下の小規模宿泊施設に分類される宿泊施設（ペンションや民宿）についてのイールドマネジメントについて検討.

小規模宿泊施設においては、価格変更のための市場調査のリソースを持たない事業者も多く、ブランド化等の検討なしに創業以来価格を変更していない事業者に対して、**Web**データを用いたイールドマネジメント手法を提案.

課題

地域の宿泊施設（ペンション等）は財を供給するにあたり、しばしば、それが**1**室（多くとも**8,9** 室など）といった限定的な数量である. このため動的な価格決定に際して、顧客セグメント別に価格調整を行うことが難しい.

宿泊施設の規模（部屋数）の違いによるデータの差異

■ 1部屋のみの場合

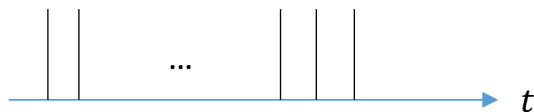


×



difficult

■ 複数部屋を持つ場合



○

予約の発生

予約数の推移

トレンド，ボラティリティ
の推定



トレンド，ボラティリティに応じた動的価格付けが可能

Webデータにもとづく類似度を用いた小規模宿泊施設の イールドマネジメント手法

宿泊施設の**Web**予約サービス等を参照することで、宿泊施設の稼働状況や売れ行きに関する分析が可能となっている。また宿泊施設についての特徴量は、宿泊施設紹介として客室のタイプ、外観・内観、立地、アメニティと詳細に掲載されており、そうしたデータが活用できるようになっている。限定的な供給量を持つ宿泊施設について、それに類似する施設を単一の仮想的な宿泊施設と見做すことで、イールドマネジメントを検討する。

提案手法

1. 対象の宿泊施設を含む観測可能な宿泊施設を特徴空間上に写像し、対象の宿泊施設との類似度が c 以上の宿泊施設について、総客室数が n を初めて超えるまで（さもなくば、類似度が c 以上の宿泊施設について）宿泊施設を選択し、この宿泊施設の集合を A とする。
2. この A について、これに含まれる全ての宿泊施設の客室および宿泊プランを用いて、仮想的な客室数 n の宿泊施設 H_A とする。
3. 仮想的な宿泊施設 H_A について、次に示す定式化を考える。

イールドマネジメントのための定式化

定式化

1. 宿泊施設 H_A がもつ全てのプランの集合を $N = \{1, 2, \dots, n\}$ とする. ここで i 番目のプランの価格を r_i とする. 仮定: i 番目のクラスの客は, i 番目のプランが売り出されていればこれを予約, そうでなければ何もしないものとする.
2. 計画期間を T 期前の予約から考えるものとし, 宿泊日を 0 期とする. 各期においては, 予約が重複して行われないものとする. e.g. Webサービス予約によるトランザクション管理
3. 第 t 期目においてクラス i の客が予約を行う確率を $p_{i,t}$ とする. $p_{0,t}$ を予約が行われない確率とする.
4. 第 t 期において x 室が既に予約済みである場合, 第 t 期から 0 期までの期待収入 $V_t(x)$ は, 以下の最適性を満たす.

$$V_t(x) = \sum_{i=1}^m p_{i,t} \max\{r_i + V_{t-1}(x+1), V_{t-1}(x)\} + p_{0,t} V_{t-1}(x)$$

ただし $t = 1, \dots, T$, $x = 0, 1, \dots, n$ であり, $V_0(x) = -R \max(0, x - n)$, R は十分大きい数とする.

5. 1プランの期待価値は $V_t(x) - V_{t-1}(x+1)$ で定まり, 第 t 期においてクラス i の客が予約を行う場合の最適戦略の条件は, r_i が1プランの期待価値を上回ることである.

簡単に！

富山県：74軒が見つかりました

[当サイト厳選](#) [料金が安い順](#) [ホテルランク](#) [ゲストの評価](#)

 マップ表示



アパヴィラホテル富山駅前 ★★★
富山市
2名が閲覧中です
本日11件の予約がありました

ダブルルーム
お得な料金で予約！
キャンセルもOK
残りあと5室です！
キャンセル無料
&前払い不要

¥13,000
-19% ¥16,000-
お部屋を選択



ホテルグランテラス富山 ★★★★★
富山市
クレジットカードなしで予約可能
本日7件の予約がありました

ツインルーム
残りあと3室です！

¥23,306
お部屋を選択



延対寺荘 ★★★
黒部市

とても良い 8.1
クチコミ82件

類似度の高いホテルを
ある客室数のもとで一つの
ホテルと見做す

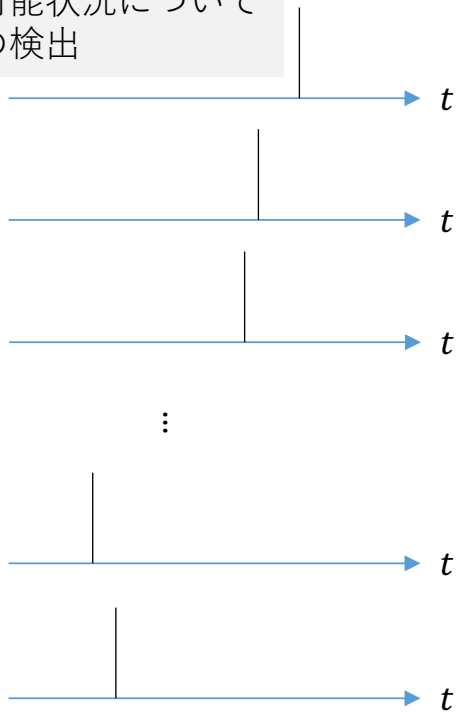
部屋タイプ	定員	本日の料金	予約条件	部屋数を選択	予約
ダブルルーム 喫煙 ベッド： ダブルベッド1台 エアコン バスタブ 専用バスルーム 薄型テレビ 無料WiFi 詳細	2人	1泊2名の合計 ¥16,000- ¥13,000 込：消費税/VAT 8%，サービス料 10%	• 2016年10月23 日までキャンセル 料無料 • 滞在時に宿泊施 設へお支払いを ください・前払い 不要 • 食事なし	0 5 残りあと5室で す！	予約画面へ 予約手数料もク レジット手数料も不 要！ 3名が閲覧中です
シングルルーム 喫煙 ベッド： ダブルベッド1台 エアコン バスタブ 専用バスルーム 薄型テレビ 無料WiFi 詳細	1人	1泊1名の合計 ¥9,500 込：消費税/VAT 8%，サービス料 10%	• 2016年10月23 日までキャンセル 料無料 • 滞在時に宿泊施 設へお支払いを ください・前払い 不要 • 食事なし	0 4 残りあと4室で す！	
シングルルーム 朝食付 喫煙 ベッド：シングルベッド1台 エアコン バスタブ 専用バスルーム 薄型テレビ 無料WiFi 詳細	1人	1泊1名の合計 ¥10,600 込：消費税/VAT 8%，サービス料 10%，朝食	• 2016年10月23 日までキャンセル 料無料 • 滞在時に宿泊施 設へお支払いを ください・前払い 不要 • 朝食込	0 4 残りあと4室で す！	
ダブルルーム 朝食付 喫煙 ベッド： ダブルベッド1台 エアコン バスタブ 専用バスルーム 薄型テレビ 無料WiFi 詳細	2人	1泊2名の合計 ¥15,200 込：消費税/VAT 8%，サービス料 10%，朝食	• 2016年10月23 日までキャンセル 料無料 • 滞在時に宿泊施 設へお支払いを ください・前払い 不要 • 朝食込	0 5 残りあと5室で す！	

仮想的なホテルの全てのプラン
についてWebセンサライズにより
予約状況を観察し、 x の推移とする

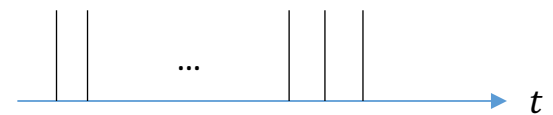
> 変動費が小さく需要の予測に依存する

仮想的な宿泊施設の予約推移

予約可能状況について
差分の検出

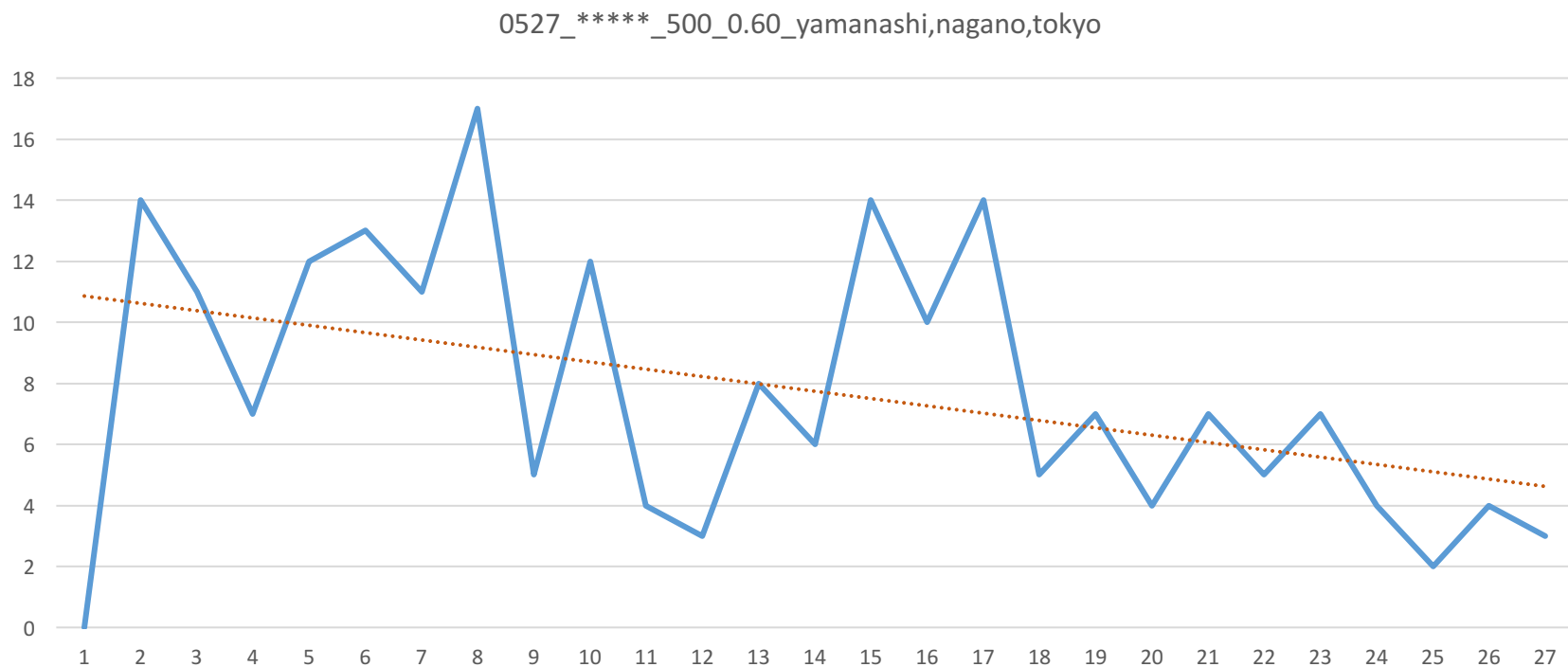


類似度 c による
宿泊施設の
マージ



総部屋数に占める時刻 t での予約数から $p_{i,t}$

+ 差が生じるプランの
価格 $p_{i,t}$ についての情報

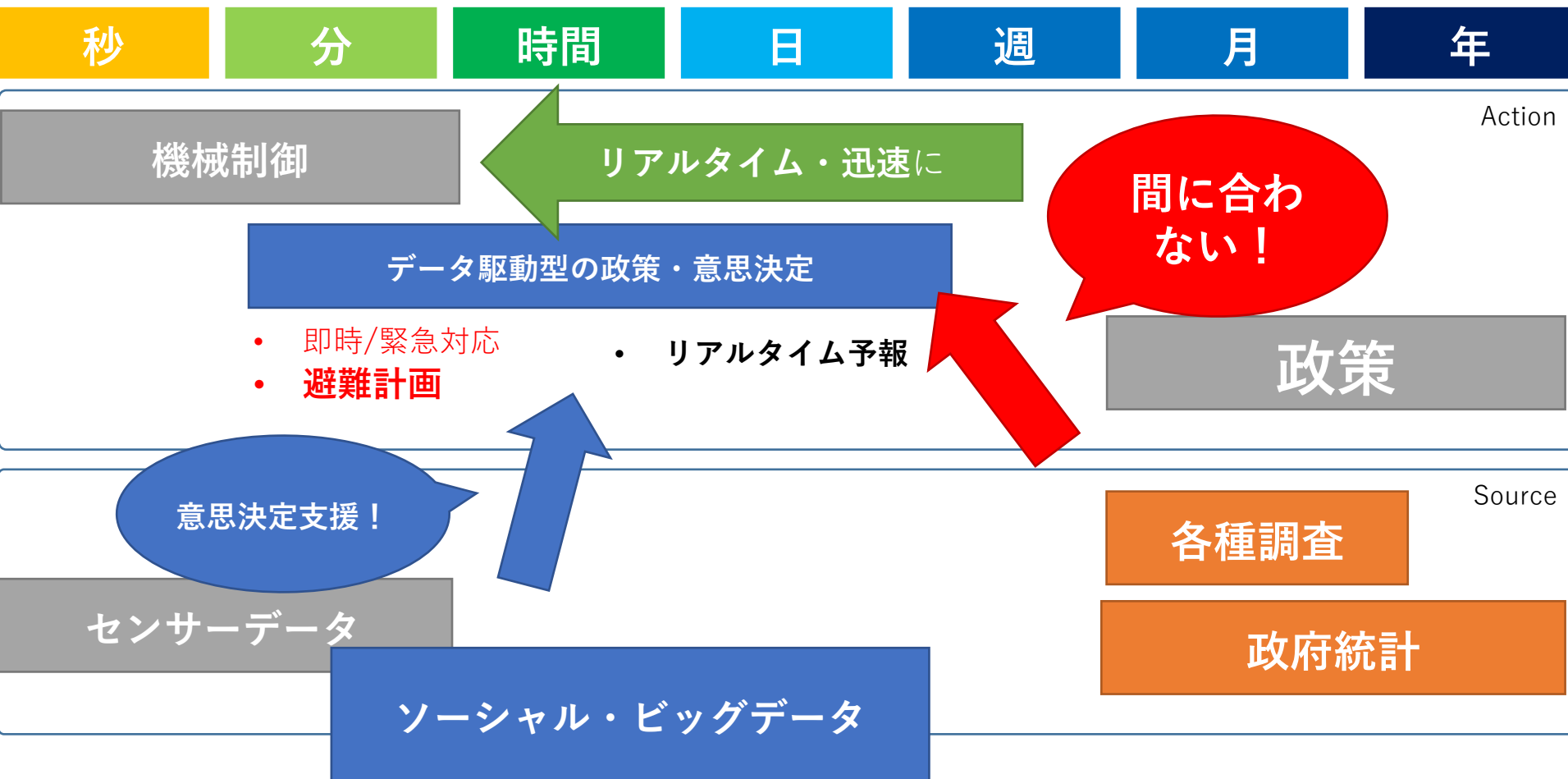


例. 5月27日の類似度0.6以上, 客室数500で形成される仮想的な宿泊施設



この推移をもとに価格設定を考えることが可能 !

意思決定の時間間隔 - 年次・月次から日次・時間単位へ



本日の講演内容

1. Web/Wi-Fiに関するビッグデータ
2. 現実世界の様子を捉え
3. どのように活用するかについて

① データを集める技術

② 分析の方法

③ 最適化への活用

見える化することで意思決定が出来るように!!

まとめ

- データの取得についてWebとWi-Fiを紹介
 - (Wi-Fi) Association, MACアドレス
 - (Web) スクレイピングとクロール
- Web/Wi-Fiによって統計データよりも早く、宿泊施設の稼働率や観光客の数を知ることが可能
 - リアルタイム
- データの活用について研究例を紹介
 - 消防署管轄についての最適化
 - 宿泊施設の予約傾向と地理特性, 効率性評価
 - イールドマネジメント