

建物活動プロファイル：建物内活動のモデル化と学習

Building Activity Profile: a Method for Modeling Activities in a Building

高井 勇志

Takeshi Takai

株式会社 竹中工務店 技術研究所

Takenaka Corporation

This paper presents “Building Activity Profile (BAP),” which describes activities in a building by a hierarchy of behavior of facilities, including electric power, HVAC, lighting, temperature, *etc.* We presume that the behavior is modeled by a hybrid system which represents that the dynamics of continuous data is derived from a combination of certain discrete modes, and represent BAP by constructing a hierarchy of the estimated hybrid systems. In this paper, we describe a framework of BAP, and shows its performance using “big data” obtained from a real building.

1. はじめに

人が24時間生活可能な設備や機能を持つ大規模ビルはいわば人間社会の縮図であり、そこで行われている活動を理解し、適切に管理・制御することは、単に一つの建物にとどまらず、地域社会全体の持続的な発展の鍵となる。

ビル由来のデータ分析によるエネルギーマネジメントについては、日本のみならず世界各国で研究開発が行われており、特に、建物で使用するエネルギーと生成するエネルギーの総和を0にすることを旨とする ZEB (zero-energy building) の実現は社会的課題となっている [NEDO]。

このような課題を解決するために、ビルの消費電力や空調、温度など様々な要素を計測できる環境が整いつつある [Kasuya 15] が、我々は、このデータ収集環境を単なるエネルギーマネジメントなどに用いるのではなく、建物内での様々な活動を理解し、建物や街に新たなサービスを提供するための基盤技術として、『建物活動プロファイル：ビルプロ』を提案する。このビルプロでは、1) ビルに設置された個々のセンサデータについて、その連続的な量の変化が何らかの離散的なモードに依存して生成されるとしてモデル化し、2) これらのモデルを階層的に扱うことによって、ビル全体の活動を表現できる。

以下では、まず、ビルのデータ通信基盤として開発しているビルコミュニケーションシステムを紹介し、その後、ビルプロの数理モデルとその学習アルゴリズムについて述べる。評価実験では、実際にビルコミを導入している施設の2年分のデータを用いてビルプロを生成し、建物における活動の分析を行うことによってその有用性を示す。

2. 建物におけるデータ通信基盤：ビルコミ

従来の建物内の設備システムでは、閉鎖的なネットワーク構成や独自プロトコルが用いられてきたが、インターネットプロトコルによる設備ネットワークの整備が進み、BACnet[BACnet] や LonWorks[LonWorks] などのオープンプロトコルが広く使われ始めている。このような背景を受けて、システムの柔軟性、リアルタイム性、初期導入コストの低減、セキュリティリスク低減などを実現するデータ通信基盤として、我々はビルコミュニケーションシステム[®] (ビルコミ)を開発している

[Kasuya 14]。

ビルコミを用いると、電力や温度などの計測・計量データのみならず、設備機器の制御情報など、ビルシステムに流れるすべての情報を集めることができる。さらに、スマートフォンやウェアラブル機器などをビルコミに接続することによって、人のデータについてもデータを収集することが可能となり¹、また、ビル内で活動するロボットからの情報もビルコミに載せることによって、ビルで行われるすべての活動についてのデータを一括して収集することが可能となる。

3. 建物活動プロファイル：ビルプロ

3.1 モデル定義

本システムで扱う入力データは、設備機器の電力値のような連続的に変化する時系列データである。この時系列データが離散的に定義できるモード系列から出現するとし、これらのモード遷移と時系列データ生成パラメータとによって定義するハイブリッドシステムを用いてビルプロのモデル化を行う。

3.1.1 単一センサデータ (図 1)

センサから得られた計測データ x_t は ($t = 1, 2, \dots$) と等間隔時でサンプリングされているとする。この時系列データを適当な区間 $X_i = \{x_{t_a}, \dots, x_{t_b}\}$ に区切り (以下、区分化時系列データと呼ぶ)、その区間に適切なモード $Q_i = q_j$ ($j = 1, \dots, M$) を割り当てることによって、単一センサについてのモデル化を行う (モードの推定方法については3.2節を参照)。ただし、 M はモードの数である。この区分化の方法についてはハイブリッドシステム設計の大きな課題の1つであるが、本稿では、多くのビルでの活動は日単位で行われていることから1日ごとに区分化を行う²。

以上により、1つのセンサについて、区分化時系列データとそれに対応するモードの組み合わせ (X_i, Q_i) ($i = 1, \dots, D$) の系列が得られることになる。ここで、 D はデータ計測の日数である。この得られたモード系列について、そのモード遷移を確率状態オートマトンで表現する。すなわち、モード q_m から q_n へ遷移する確率を $P(q_n|q_m)$ と表現する。また、各モードか

*1 データ収集に対する同意を得ることや、プライバシーやセキュリティなどの問題は当然解決する必要がある。

*2 交代制の工場など日付に依らない区分化もあり得るため、データに基づいた適切な区分化法が必要ではあるが、これについては今後の課題としたい。

ら出現するダイナミクス $\hat{X}_j$ を、平均 μ_j 、分散共分散行列 Σ_j で表される正規分布 $N(\mu_j, \Sigma_j)$ ($j = 1, \dots, M$) で表現する。

3.1.2 センサ集合データ：複数センサの統合 (図 2)

建物には前節で述べたようなセンサが複数設置されているが、これらを統合して扱い、センサ集合のモードとその変化によって建物全体の活動を表現する。

前節の記法を用いると、建物に S 個のセンサがある場合、ある 1 日のセンサ集合のモード組は $(Q_i^{(1)}, Q_i^{(2)}, \dots, Q_i^{(S)}) \equiv Q_i$ と書ける。この Q_i に適切なモード $Z_i = z_k$ ($k = 1, \dots, N$) を割り当てると、センサ集合のモードを表現することができる。つまり、センサ集合についても、モード組とそれに対応するモードの組み合わせ (Q_i, Z_i) ($i = 1, \dots, D$) の系列が得られることになる。またモード遷移についても単一センサの場合と同様に表現し、モード z_m から z_n へ遷移する確率を $P(z_n|z_m)$ と書くこととする。

3.2 学習アルゴリズム

本節では、まず、単一センサにおけるモード推定やその遷移確率、また、推定したモードから生成されるダイナミクスの生成モデルの学習について述べる。その後、単一センサのモデルを統合することによるセンサ集合のモード学習について述べる。

3.3 単一センサ (図 3)

各センサのモード学習は以下の 3 つのステップにて行う。

1. 次元削減 各センサの区分時系列データセット $X = \{X_1, \dots, X_D\}$ に対して主成分分析を行い、次元削減を行ったデータセット $X' = \{X'_1, \dots, X'_D\}$ を得る。ここで、 X と X' の同じ添字の要素は同じ日のデータを指すものとし、縮約する次元数は累積寄与率を用いて決定する。

2. クラスタリング 得られた X' に対して、混合正規分布推定によるクラスタリングを行う。正規分布の数 (つまりクラスタ数) については、予め指定した範囲で数を変えながら混合正規分布推定を行い、バイズ情報量基準が最小となる数を選択する。ここで得られたクラスタ数がそのセンサデータに対して出現するモードの数となり、モードの集合 $Q = \{q_1, \dots, q_M\}$ と各モードに属するデータのサブセット $X'(q_j) \subseteq X'$ ($j = 1, \dots, M$) が得られる。

得られたモード q_j のダイナミクス生成パラメータ μ_j, Σ_j は、 $X'(q_j)$ に対応する区分時系列データの $X(q_j) \subseteq X$ を用いて、この平均と分散共分散行列を求めることによって得ることができる。ただし、 Σ_j については対角成分のみを用いることとしその他の成分は 0 とする³。

3. モード遷移確率計算 モードの遷移確率 $P(q_n|q_m)$ を以下の

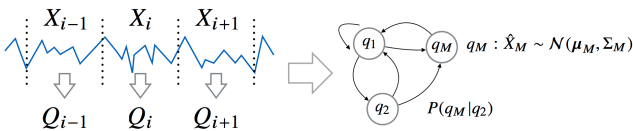


図 1: 単一センサデータのモデル化。

*3 データの次元数に比べてサンプル数が少ない場合は Σ の逆行列が得られず、後述する KLD が求められない場合があるため。

ように求める。

$$P(q_n|q_m) = \frac{P(q_n, q_m)}{P(q_m)} = \frac{N_{q_n, q_m}}{N_{q_m}}. \quad (1)$$

ここで、 N_{q_n, q_m} は X の中で連続する区分時系列データ X_i, X_{i+1} のモードがそれぞれ q_m, q_n となる数、 N_{q_m} は $X(q_m)$ の数である。

以上によって、単一センサのモデルの学習が行われる。

3.4 センサ集合：複数センサの統合 (図 4)

複数センサの統合についても、単一センサの場合と同様の考え方でモデル化を行う。つまり、センサ集合の活動もいくつかのモードに分けられると仮定すると、すべてのモード組 Q をある適当な特徴空間に投影してクラスタリングを行うことによって、モード組のモード Z を推定することができる。また、そのモード間の遷移確率も、単一センサの場合と同様にして求めることができる。

ただし、モード組 Q はモードを指す質的変数であるため、単純に主成分分析を行うことはできない。そこで、モード間の差異を定義し、差異を基に分布を求める方法である多次元尺度構成法 (MDS) を用いて、質的変数であるモード q から量的変数であるモード値 d に変換し、モード組 Q に対応するモード値の組 $\mathcal{D}$ を得る。これを用いてクラスタリングを行った後に、モードの遷移確率を求めることになる。

以上をまとめると下記のようなアルゴリズムとなる。

1. 各センサのモード差異行列の取得 モード間の差異は、各モードが生成するダイナミクス $\hat{X}$ を表す正規分布のパラメータを用いてカルバック・ライブラー情報量 (KLD) によって求める。ただし、KLD には対称性がないため、モードの差異を

$$\delta(q_a, q_b) = \frac{D_{KL}(q_a||q_b) + D_{KL}(q_b||q_a)}{2}. \quad (2)$$

と表す。ここで、 $D_{KL}(q_a||q_b)$ はモード q_a とモード q_b の KLD である。各センサについて、モード間の差異を式 (2) を用いて求めると、これを要素とするモード差異行列 $M^{(s)}$ が得られる。

$$M_{ij}^{(s)} \equiv \delta(q_i^{(s)}, q_j^{(s)}) \quad (3)$$

ここで、 $M_{ij}^{(s)}$ は $M^{(s)}$ の i 行 j 列成分である。

2. モードからモード値への変換 上記にて求めた $M^{(s)}$ を用いて、1 次元の多次元尺度構成法を行い、各モードに対応するモード値を求める。

$$d_m^{(s)} \leftarrow f_{MDS}^{(s)}(q_m^{(s)}) \quad (m = 1, \dots, M) \quad (4)$$

この式により、モード組 Q_i をモード値の組 $\mathcal{D}_i$ に変換することができる。

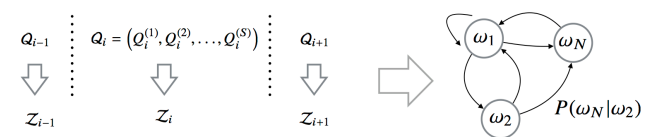


図 2: センサ集合データのモデル化。

3. 次元削減 単一センサの場合と同様に、データセット $D = \{D_1, \dots, D_D\}$ に対して主成分分析を用いた次元削減を行い、 D' を得る。ただし、 D には予め白色化を施しておく。また、 D と D' の各要素は対応が取れているものとし、縮約する次元数は累積寄与率を用いて決定する。

4. クラスタリング及びモード遷移確率計算 単一センサの場合と同様にクラスタリングを行い、モードの集合 $\Omega = \{\omega_1, \dots, \omega_N\}$ と各モードに属する $D'(\omega_k) \subseteq D'$ ($k = 1, \dots, N$) を得る。ただし、 N は得られたモードの数である。また、モードの遷移確率 $P(\omega_n|\omega_m)$ も単一センサの場合と同様に求められる。

以上により、各センサのモードを統合したセンサ集合のモード学習を行うことができる。

4. 評価実験

竹中工務店技術研究所の2014年1月1日から2015年12月31日までの2年分のデータを用いて実験を行った。データセットは施設A-Fの6箇所の消費電力、太陽光発電、風力発電、外気温、外気湿度について毎分計測した値であるが、停電作業や欠損などのため合計698日分のデータとなっている。なお、この施設A-Fの消費電力の合計が研究所全体の消費電力とほぼ一致している。

始めに、各施設等それぞれの計測データから学習したモードを図5に示す。それぞれの元データが示すように、施設によって電力消費のパターンが大きく異なっていることがわかる。施設A, B, F (図(1)–(3), (4)–(6), (16)–(18)) については日による変化(横軸)と時刻による周期性(縦軸)が見られるが、施設D (図(10)–(12)) については、日や時刻によらずほぼ一様な電力消費がなされていることがわかる。これらの推定されたモードのダイナミクスを見ると、 Σ の範囲が比較的狭く、同じモードと推定された日々での消費電力の変動の分散が少ないことがわかる。また、施設E (図(13)–(15)) を見ると、特徴的な電力消費パターンから機械的な制御が行われていることがわかり、推定されたモードのダイナミクスからは時刻による消費電力パターンの特徴によって幾つかのモードに分かれていることがわかる。施設C (図(7)–(9)) については、特定の日や時刻にのみ電力が消費されていることがわかり、これらの推定されたモードのダイナミクスの分散が比較的大きいことから、年間を通じて日々の変動が大きいくことがわかる。

太陽光発電 (図(19)–(21)) については、施設の運用上午前8時から日没までの発電が行われており、季節による大きな変動と、日照変化による発電量の違いが見取れる。また、風力発電 (図(22)–(24)) については、各モードでの Σ が大きいことから、日々や時刻での差が大きいくことがわかる。さらに、気温と湿度 (図(25)–(27), (28)–(30)) については、自然現象であるため元データが滑らかに変化していることがわかり、その変化のパターンに応じてモードが推定されていることがわかる。

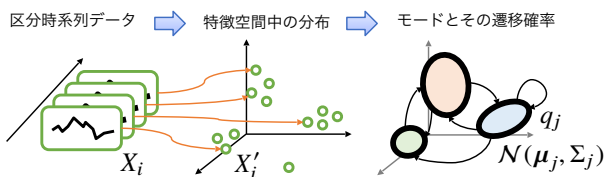


図3: 単一センサのモード学習の流れ。

また、図6に示すヒストグラムには各センサの特徴が表れているが、属する日が少ないモードが多く現れているセンサ(施設C)については、年間を通じて現れる共通のダイナミクスのパターンが少ないことを示しており、このようなセンサに対しては本手法を当てはめることが適切でない可能性がある。さらに、他のモードと比べて極端に属する日が少ないモードについては、他の差異の少ないモードと統合することによって、モード数が過剰とならないようにする処理が必要であると考えられる。

最後に、図7にセンサ集合の学習結果を示す。クラスタリングの結果、センサ集合は15個のモードに分けられた。図(3)にあるように土日や休日などに同一のモードが多く現れていることや、また、同じモードとなる日が局所的に現れていることがわかる。これは、施設の消費電力や気温などのモードから意味のあるモード組のクラスタリングが行われた結果だと考えられる。センサ集合の階層化や、モード組における各センサに対する重み付けなどを行うと、より高度なモード学習が行われると考えられるが、これについては今後の課題としたい。

5. まとめ

本稿では、建物内での活動を理解するための基盤技術として、建物活動プロフィール: ビルプロを提案した。ビルプロの技術を用いることによって、膨大な数値の羅列であるビル由来データを直感的に理解しやすい形で表現し、これを建物や街に新たなサービスを提供するための基礎情報として利用することができる。

今後日本が迎える人口減少社会においては、地方の市区町村人口が激減し、各地に点在する複数の中核都市に人口が集中する、コンパクトシティ化が進むと考えられる。つまり、1つの大規模ビルはそれだけで1つの社会となり、ビル環境が果たすべき役割はより大きくなるだろう。ビルが持つ情報は、今後ますます多種多様となりその量も増大し続け、近い将来に、人がこのような膨大な情報を適切に管理し制御することができなくなるのは明らかである。ビルプロを用いてビル内での活動を表現することによって、将来、建築空間において人工知能やロボットとの共生が行われ、持続可能な建物や街づくりへ発展することが期待できる。

参考文献

- [BACnet] BACnet. <http://www.bacnet.org>.
- [Kasuya 14] 粕谷ら. “スマートシティのための MQTT プラットフォームの検証.” FIT2014 (第13回情報科学技術フォーラム) (2014).
- [Kasuya 15] 粕谷ら. “ビッグデータリアルタイム解析による建物管理システム,” 空気調和・衛生工学 第89巻 第12号 pp.65–pp.70 (2015).
- [LonWorks] LonMark International. <http://www.lonmark.org>.
- [NEDO] 国立研究開発法人新エネルギー・産業技術総合開発機構. “ZEB (ネット・ゼロ・エネルギービル)、省エネルギービル,” http://www.nedo.go.jp/activities/DA_outline-zeb.html.

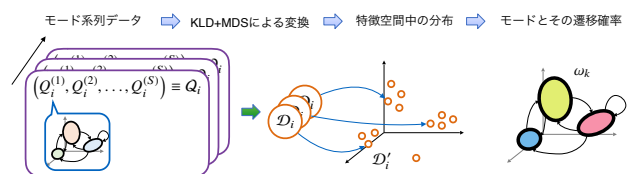


図4: センサ集合のモード学習の流れ。

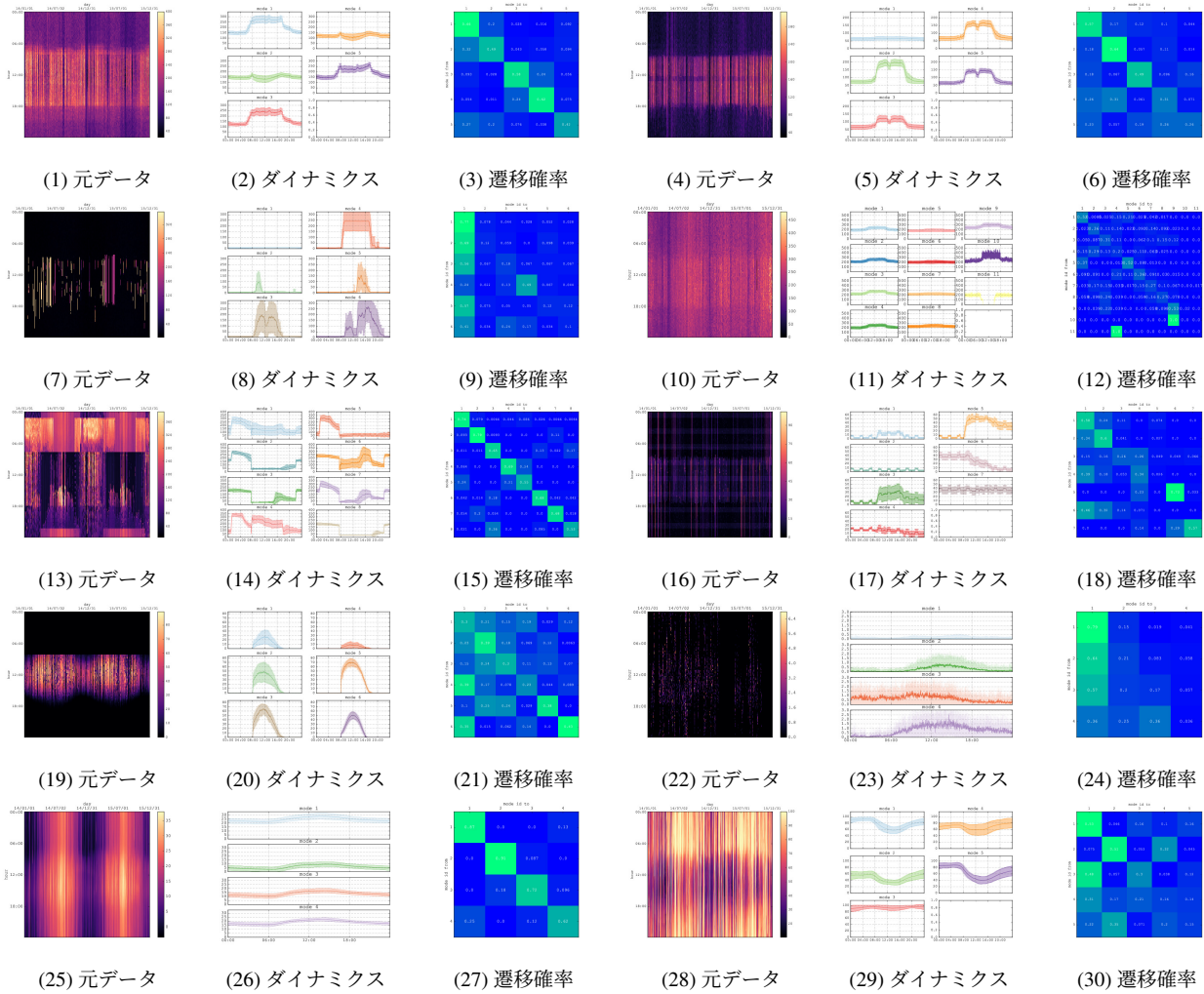


図5: 各センサの2年分のデータと学習結果。“元データ”はセンサから得られた値を, “ダイナミクス”は学習したモードの μ と Σ を, “遷移確率”はモードの遷移確率を可視化したものを示している。“元データ”の横軸は日, 縦軸は時刻であり, 明るい色ほど値が大きいことを示している。“ダイナミクス”の横軸は時刻, 縦軸は値であり, それぞれ学習したモードのパラメータである μ と Σ を, 実線とその上下範囲で示している。“遷移確率”は行列を可視化したもので, 行・列がそれぞれ遷移元・遷移先モードの遷移確率を示しており, 青から緑になるほど高い確率であることを表している。各図とデータの対応は以下の通りである。(1)―(3)は施設A, (4)―(6)は施設B, (7)―(9)は施設C, (10)―(12)は施設D, (13)―(15)は施設E, (16)―(18)は施設Fの消費電力である。また, (19)―(21)は太陽光発電, (22)―(24)は風力発電, (25)―(27)は外気気温, (28)―(30)は外気湿度である。

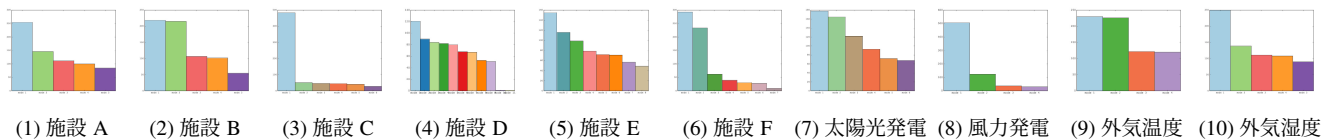


図6: 各センサのモードのヒストグラム。横軸がモード番号, 縦軸がそれぞれのモードに分類されたデータ数(日数)を示している。

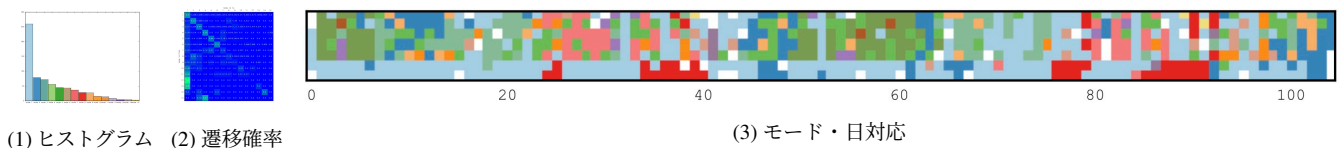


図7: センサ集合の学習結果。“モード・日対応”の縦軸は曜日(上から月, 火, …, 日), 横軸は週番号(2014年1月1日から2015年12月31日まで)を示している。また, データが欠損している日は白色で表示している。