

消費電力データと宅内センサーデータの解析

Analysis of Power Consumption Data and Indoor Sensor Data

○大桃 和樹(会津大学) 金田 祐也(会津大学)
 多屋 優人(株式会社 KDDI 研究所) 趙 強福(会津大学)

Kazuki OMOMO, University of Aizu, m5181112@u-aizu.ac.jp
 Yuya KANEDA, University of Aizu
 Masato TAYA, KDDI R&D Laboratories
 Qiangfu ZHAO, University of Aizu

We want to establish a technology to estimate household structure or at-home situation, by analysis of power consumption data, indoor sensor data, and at-home situation data. If it was established, We would provide information related to saving power to the residents, or provide services for elderly persons. The first step of this research, here we confirm whether household structure can be classified based on only power consumption data. If it is possible, device costs are reduced.

Key Words: Indoor Sensor, Power Consumption, k-means, Adjusted Rand Index, Self Organizing Feature Map

1 緒言

近年、日本においては高齢者が増加し、日本の人口に対する高齢者の割合が増加している。高齢者の中には、家族で住んでいる人がいれば、家に一人で暮らしている事もある。一人暮らしの老人の中で問題となっているのが、孤独死である。文献 [1] によれば、平成 20 年度には単身居住者で死亡から相当期間経過後に発見された件数が 154 件だったのに対し、平成 24 年度には 220 件に増加している。団塊世代の高齢化により、この問題はこれからさらに大きな問題と発展する可能性が高い。

我々は、センサーや宅内状況情報分析による、在宅状況及び世帯構成情報等を推定する技術について研究している。我々が対象とする情報は、宅内センサー、消費電力情報、及び宅内状況情報等である。本技術が確立すれば、孤独死の早期発見、または異常を検知して自動的に救急車を呼ぶ等の方法を通して、孤独死を回避する事が可能であると考えている。さらには、世帯構成及び在宅状況推定が出来れば、効率的な電力消費を促す事が可能であり、資源の節約に貢献する事が可能となる。

本稿では、消費電力データに着目して、世帯構成及び在宅状況を検知する事が可能か検証する。もし消費電力のみで検知可能であれば、導入コストを格段に抑える事ができ、より多くの家庭で導入を行う事が出来るようになる。また、検知方法として移動平均法とクラスタリング手法である k 平均法 (k-means) を基にした 2 つの手法を提案し、判別可能かを確認する。後者については、自己組織化写像 (SOFM: Self Organizing Feature Map) を用いて各クラスターを 2 次元空間へ写像し、特徴について考察する。

本稿の構成は、次の通りである。2 章では、本稿で提案する手法の関連手法について簡潔に紹介する。3 章では、今回提案する移動平均法とクラスタリング手法を基にした消費電力からの推定方法について説明する。4 章では、提案した手法について実験を行い、結果について考察する。5 章は本稿のまとめであり、今後の活動について 6 章で紹介する。

2 基礎知識

2.1 k 平均法

k 平均法 (k-means) は、非階層型クラスタリング手法の 1 つである。与えられた学習データに対し、任意の k 個のクラスターに分ける事を目的とする。その為に、各クラスターの中心値を計算し、データを判別する。

データ判別の基準となる k 個の中心値を計算する流れを、以下に示す。

1. 与えられた学習データの各データに対し、ランダムにクラスターを割り当てる。
2. 各クラスターの中心値(重心)を計算する。
3. 各データを、最も近い中心地のクラスターで割り当て直す。
4. 上記 2,3 を、全てのデータにおいてクラスター再配置が行われなくなるまで繰り返す。

2.2 調整ランド指数

ランド指数とは、2 つのクラスタリング結果の類似度を測る尺度として用いられる。しかし、2 つのクラスタリング結果の間に相関性がない場合、値が一定にならない為、値の解釈が難しいという問題が発生する。調整ランド指数 (ARI: Adjusted Rand Index) は、相関性がない場合に値が 0 となるように計算式を調整したものである [3]。ARI は式 (1) で求められる。ここで、 $Index$ は 2 つのクラスタリング結果双方において同じクラスター、または違うクラスターのペア数、 $MaxIndex$ は全ての要素のペアの数であり、 $ExpectedIndex$ は $Index$ の期待値である。

$$ARI = \frac{Index - ExpectedIndex}{MaxIndex - ExpectedIndex} \quad (1)$$

2.3 自己組織化写像

自己組織化写像 (SOFM: Self-Organizing Feature Map) とは、ニューラルネットワークの一種であり、教師なし学習を通して入力データを任意の次元へ写像することができる。本稿における写像先

は、2次元空間とする．このSOFMでは2次元上にノードが配置され(Fig. 1 参照)，各ノードは入力データと同次元の重みベクトルが対応付けられている．入力データに最も近い重みベクトルを勝者とし，勝者とその近傍の重みベクトルを入力データに近づけることで学習する．この動作を様々な入力に対して繰り返し行うことで，入力データ同士の関係性を保存した写像を得ることが出来る．

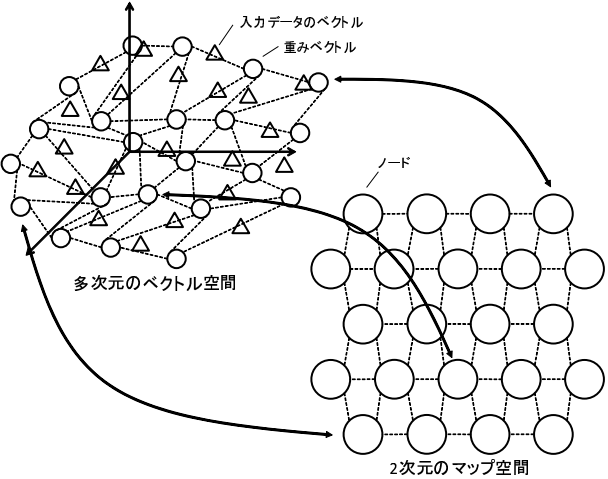


Fig.1 Figure of SOFM

2.4 移動平均

移動平均 (MA: Moving Average) は、時系列データを分析する際に用いられる計算法の一種である．本稿では、単純移動平均を用いる．単純移動平均とは、与えられた N 個のデータの集合 $U = \{a_1, a_2, \dots, a_N\}$ から、直近の $n(n < N)$ 個のデータを用いて重み付け等を行わず平均値 MA_i を計算する．計算式を、式 (2) に示す．集合 U から得られる単純移動平均値は、 $N - n + 1$ 個となる．

$$MA_i = \frac{1}{n} \sum_{j=0}^n a_{i+j}, \text{ for } i = 1, 2, \dots, N - n + 1 \quad (2)$$

3 消費電力データを用いた世帯構成及び在宅状況推定法

本稿では、世帯構成及び在宅状況を推定する為に、消費電力データを対象とした2つの手法を提案する．1つは移動平均を用いた行動予測についてであり、もう一方は、クラスタリングによる世帯構成推定である．

3.1 移動平均を用いた在宅状況推定法

消費電力データから住民の在宅を推定する方法として、我々は移動平均を用いた手法を考えた．計測された各日の消費電力積算値だけではノイズが大きく、行動推定が難しいと判断したからである．そこで、我々は D 日移動平均 (D -day MA) を計算し、消費電力積算値と合わせて行動推測をする方法を考えた．計算値を確認する為に、推定期間におけるグラフを準備し、行動推定を行う．

3.2 クラスタリングによる世帯構成推定法

消費電力データから家庭の世帯構成を推定する為に、我々はクラスタリング手法を基にした方法を考えた．本稿で用いた世帯情報は、Table 1 の通りである．

Table 1 Feature of household

項目	区分	単位
居住者数	大人, 子供	人
世帯情報	単世帯, 複数世帯	世帯
延べ床面積		m^2
暖房機器	電気, ガス, 灯油	
住宅形態	戸建, マンション, アパート	
生活形態	定住, 一時滞在	

本稿では、クラスタリングの手法として、非階層型クラスタリング手法として知られる k-means を用いた．また、各クラスタリング結果の特徴を確認する為に、クラスタリング結果について SOFM を用いて2次元空間へ写像を行い、写像化された結果を基に推定を行う．

4 実験

本稿では、提案した方法によって推定が可能かどうか、実験を通して考察を行う．実験内容は2種類であり、移動平均を用いた在宅状況推定についてと、クラスタリング手法を基にした世帯推定についてである．

4.1 移動平均を用いた在宅状況推定について

我々は、まず移動平均を用いた在宅状況推定について確認する為、観測データを用いて実験を行った．

4.1.1 実験設定

今回用いた観測データは、会津若松スマートシティ推進協議会から提供されたデータを用いた．データの収集には、Fig. 2 にあるスマートメーターを用いた．対象とした世帯は3世帯であり、それぞれ異なる期間を対象としている．それぞれの期間を、Table 2 に示す．

Table 2 Terms of experiment in each household.

世帯	期間
世帯 A	2013 年 11 月 13 日から 2014 年 3 月 9 日
世帯 B	2013 年 11 月 17 日から 2014 年 3 月 9 日
世帯 C	2013 年 3 月 11 日から 2014 年 3 月 9 日

本実験では各日の消費電力積算値を対象に、5 日間移動平均 (5-day MA) と 25 日間移動平均 (25-day MA) を計算した．計算された値は、各日の消費電力積算値と共に折れ線グラフで表示した．

4.1.2 実験結果と考察

Figs. 3-5 は、3つの世帯についてのグラフである．それぞれのグラフでは、グラフの傾向を読み取りやすくするために、消費電力積算値と併せて5日移動平均(5-day MA) と 25 日移動平均(25-day MA) を表示している．

Figs. 3-5 を確認すると、世帯 A, B 共に 1 月中旬ごろまで上昇し、その後は少し減少している．これは、暖房器具などによる電力の消費が、気温が寒くなるにつれて増加したためではないかと推測できる．また、世帯 B のグラフの大晦日・正月ごろの消費電力が極端に減少している事から、この家庭は大晦日・正月に実家に帰省していたのではないかと推測できる．世帯 A のグラフでは、大晦日・正月の消費電力が増加している事から、この家庭は大晦日・正月は家で過ごしていたのではないのかと推測できる．



Fig.2 Smart meter

世帯 C のグラフは、今回対象とした世帯で最も期間が長い世帯である。この結果から、冬場になると消費電力が増加している。8月10日に極端に消費電力が増加しているが、この日は猛暑日であった。よって、冷房の使用が増加したためではないかと推測する事が出来る。

以上から、消費電力データから家族全体の行動(帰省や冷暖房器具の使用など) をある程度推定出来ると言える。時間ごとの積算値を見れば、特定の時間帯にお風呂、睡眠、在宅、食事、料理などの行動も、推定出来る可能性があるため、今後は時間毎のデータにも着目して考察していきたい。

4.2 クラスタリング手法を基にした世帯推測について

本実験では、消費電力と世帯情報を対象として、クラスタリング手法を基にした世帯推測について考察を行う。

4.2.1 実験設定

本実験では、48 世帯の消費電力データと世帯情報を対象とし、それぞれにクラスタリングを行った。クラスタリングには k-means 法を使用し、分類するクラスタ数 k は 5, 10, 15 とした。

まず、消費電力データを基に世帯をクラスタリングを行った。1つの世帯を表す特徴ベクトルは、次の通りである。各時間(0~23 時) 別に 1ヶ月の消費電力の平均値を計算する。その結果、24 次元のデータが得られ、1つの世帯の特徴ベクトルとする。今回は、すべての世帯において 4.1 章の実験と同じ期間内のデータを用いた。

また、同様に世帯情報を対象としてクラスタリング実験を行った。世帯情報についての特徴ベクトルは、Table 1 で示したものを利用した。

これら 2つのクラスタリング結果が近いものであれば、消費電力データのみから世帯情報をある程度読み取り、世帯を分類できたと言える。この比較のために、ARI を用いた。

さらに、各クラスタリング結果の特徴を調べるため、SOFM を用いて各クラスタリング結果を 2次元へ写像し、その特徴を調べた。SOFM として、ノードを 11×11 の六角格子状に並べたものを用いた。SOFM の学習回数は 10000 回とした(1つの入力データに関する 1回のノードの重み更新を、学習回数 1回とする)。

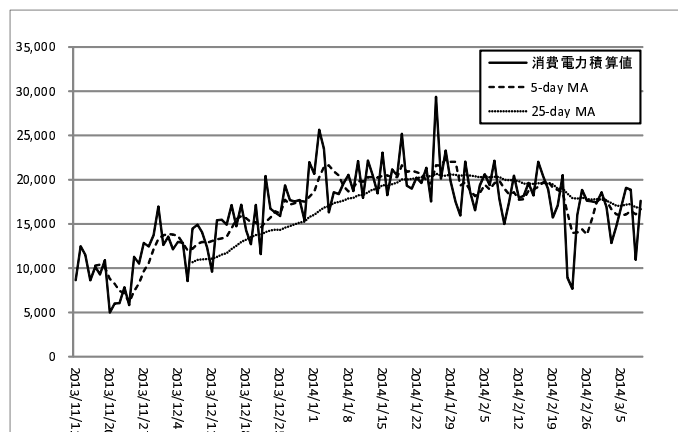


Fig.3 Power consumption of household A

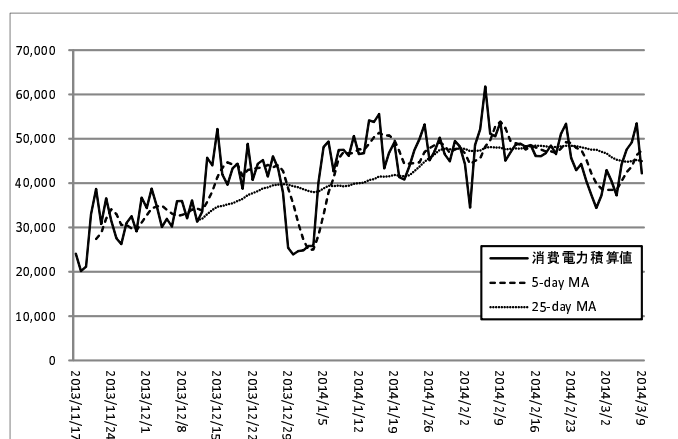


Fig.4 Power consumption of household B

クラスタ数が多いと特徴が読み取りにくいいため、 $k=5$ の場合のみ行った。

4.2.2 実験結果と考察

Table 3 Result of similarity based on ARI

k	類似度
5	52.79%
10	54.20%
15	60.72%

Table 3 は、 k の値による ARI を用いて計算された類似度の変化を表している。また、Figs. 6-9 は SOFM によって 2次元に写像された時の特徴を可視化した結果である。

k がどの値の場合でも、約 50% という結果であった。世帯構成を推測する上で、50% という数字はランダムに判定した場合でも同等の数値が出る可能性があるため、精度がまだ足りないと考えられる。 k の値が増えるほど類似度が高くなっていることについては、消費電力を基にした分類が細かいほど、正確に世帯を分類できるのではないかと考える事ができる。

次に、2つのクラスタリング結果を SOFM で可視化した図を確認する。今回は読み取りやすさのため、 $k=5$ の時のみを可視化している。各クラスタリング方法について、2回分の結果を用意した。

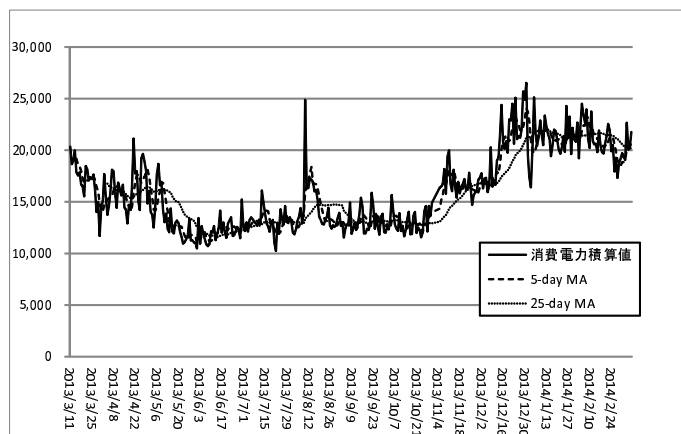


Fig.5 Power consumption of household C

Fig. 6, 7 は消費電力データを基にした世帯のクラスタリングの結果である。Fig. 8, 9 は世帯情報を基にした世帯のクラスタリングの結果である。図中の A, B, C, D, E は 5 つのクラスタを表している。Fig. 6, 7 において、各クラスタはクラスタ E を中心として放射状にマッピングされている。Fig. 8, 9 においては、各クラスタは D, E, A, B, C の順に、縞模様を描いてマッピングされている。各クラスタリング結果はそれぞれ特徴を持っていたが、2 つのクラスタリング結果は異なっていた。

Fig. 6 におけるクラスタ E に着目してみた。このクラスタに属する世帯の消費電力データについて調べてみたところ、共通する特徴を見つけた。それは、8 時から 16 時までの消費電力は低く、それ以外の時間帯は高い、という特徴である。このことから、このクラスタに属する世帯は、8 時から 16 時までには居住者は外出していて、それ以外の時間帯は家に居てかつ電力を消費していると推測できる。このケースでは、居住者の行動のみでクラスタリングされた可能性があり、消費電力データと世帯情報の間には、相関性がない場合があると考えられる。居住者が家で活動しているような時間帯（例えば 18 時～0 時）に絞ってみると、異なる結果が得られる可能性があるため、確認していきたい。

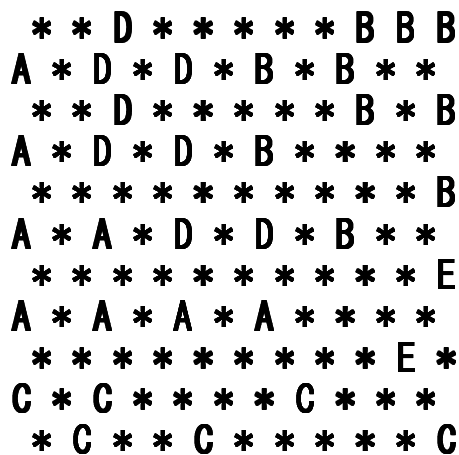


Fig.6 Visualization of a clustering result based on power consumption

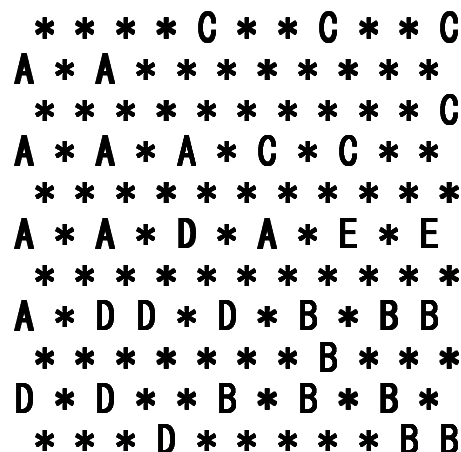


Fig.7 Visualization of a clustering result based on power consumption

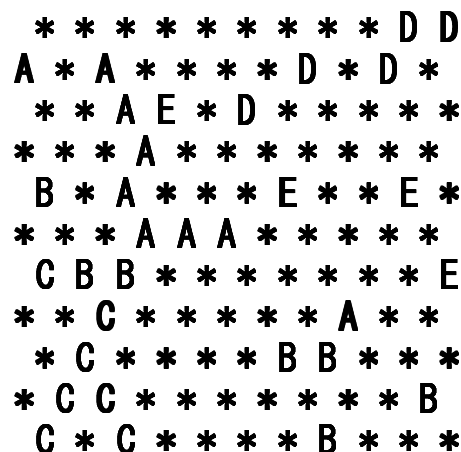


Fig.8 Visualization of a clustering result based on household information

5 結論

本稿では、消費電力、宅内センサー情報、及び在宅状況情報データ分析による、世帯構成や在宅状況を推定する技術の確立を目指す最初の段階として、消費電力データのみに着目して検証した。消費電力データのみから世帯に関する正確な情報を読み取るまでにはいかなかったが、世帯構成や在宅状況は消費電力データにある程度関連しているという事がわかった。

1 日積算消費電力値をグラフ化した検証では、消費電力データのみから、帰省や冷暖房の使用などの、居住者の行動について推定する事が出来た。消費電力データのみからでも、居住者のある程度の行動の推定が可能であると言える。

消費電力データを用いた世帯のクラスタリングでは、消費電力データから世帯情報のある程度読み取れることがわかった。また、消費電力データは、居住者の外出のタイミングに大きく依存していることに気をつける必要があることがわかった。


```

* * * * * C * C * C
B * B * C * * * C * *
* * * * C * C C C * *
* B B * * * * * * *
B * * * * * * * B * *
A * * A A * * * * B *
* * * * A * * * * * E
A * A * A * * A * * *
* * * * * E E * * * *
E * * * * * * * * D *
* E * * * * D * D * * D

```

Fig.9 Visualization of a clustering result based on household information

6 今後の活動

本稿では、消費電力データのための検証を行った。その結果を生かし、より詳細な推測を目指す研究をしたいと考えている。そのため、次に挙げる点について、これから検証していく。

1つ目として、1日積算消費電力のグラフ化において課題となった、お風呂、睡眠、在宅、食事、料理などの詳細な行動の解析である。今回は1日ごとの積算値を比較して、帰省と冷暖房器具の使用という2つの推測を挙げた。時間ごとの積算値のグラフを解析し、より詳細な居住者の行動についての解析をする必要があると考える。

2つ目として、消費電力データを用いた世帯のクラスタリングにおいて課題となった、居住者が家で活動している時間帯に絞った場合にどのような結果となるか確かめることである。絞った時間帯においては居住者全員はほぼ家にいる状態であり、消費電力の変化は居住者の不在を考慮することなく、居住者の家の中での行動に強く依存していた。この時間帯のみに着目することで、居住者の行動や世帯の情報などを読み取りやすくなり、正確な推定につながるのではないかと考えられる。

その他としては、消費電力データから世帯の正確な人数を推定する検証をしたいと考えている。消費電力データを基に特徴ベクトルを作成し、教師信号として世帯の人数を設定した上で、ニューラルネットワークなどを用いた手法が効果的である可能性がある。消費電力データから特徴ベクトルを作成する際に、本稿の検証結果や、2つの課題の結果を参考にしたい。

References

- [1] 内閣府, “平成 26 年版高齢社会白書 (全体版)”, http://www8.cao.go.jp/kourei/whitepaper/w-2014/zenbun/26pdf_index.html
- [2] T. Kohonen, “Self-Organizing Maps”, Springer, 2000.
- [3] Hubert, L., & Arabie, P., “Comparing partitions,” Journal of Classification 2 (1): 193–218, 1985.