

主幹電力時系列データからの家庭内行動推定手法

石津 紘太郎^{*1} 山口 弘純^{*1} 東野 輝夫^{*1}^{*1} 大阪大学 大学院情報科学研究科 {k-ishidu, h-yamagu, higashino}@ist.osaka-u.ac.jp

1 はじめに

近年、センシング技術の目覚ましい発展に伴い、家庭内の居住者の行動情報をセンサ等で認識し、様々なサービスに活用することが期待されている。それに対し、これまでに多くの行動認識技術が提案されてきている。動画像や静止画像を用いる行動推定手法 [1] は最も直接的であるが、プライバシー懸念から家庭におけるカメラの受容度は極めて低い。また、ウェアラブルセンサを用いた行動推定手法 [2] は、常時センサを装着する必要がある、装着忘れによるデータの取りこぼしや居住者の負担増が想定される。また、侵襲性の低い設置型の低解像度センサを複数利用し、プライバシーに配慮しながら行動認識を行う技術についても提案されているが、家庭内でのセンサ設置とデータ集約機構、ならびにデータ送信のための通信回線が必要となり、単身の高齢者宅等への導入の障壁となる可能性がある。

本研究では、家庭の総消費電力（主幹電力）データのみを用い、家庭内の大まかな行動を推定する手法を提案する。主幹電力はスマートメータ等から管理目的で定期的に取得されるため、新たなセンサシステム設置が不要であり、コストかつ侵襲性の点において大きな利点がある。

一般に、家電の起動時には特有の消費電力の変化が生じることが多いため、それを把握できればどういった家電が起動したかが把握でき、それに応じた行動推定も可能である [3] もの、主幹電力データは一定時間毎に稼働した家電の消費電力や待機電力に加え、給湯器・温水洗浄便座といった宅内電気設備の消費電力の総量として取得されるため、個別の家電の種別や稼働状況の把握は困難である。これに対し、提案手法では家庭内行動の多くが家電の利用を伴うことに着目し、調理時間中や睡眠時間中にみられる主幹電力の大まかな増加や減少を特徴量とし、教師あり機械学習による行動推定を行う。ここで、家庭内行動と電力消費傾向は家庭毎に異なるが、各対象家庭から教師データを収集し、家庭毎の分類器を構築することは現実的でない。また未知の対象に対し学習済分類器を転用する転移学習なども知られているが、対象家庭の少量の教師データが必要となる。そこで提案手法では、複数のモデル家庭から収集した教師データから、それらの家庭毎に分類器を構築し、対象家庭のテストデータに対し適切な分類器を選択することで対象家庭の教師データを不要とする手法を提案している。

提案手法の評価のため、実在 10 家庭の配電盤にクランプを設置したうえで 30 秒毎の主幹電力データを 191 日間取得したデータセットを利用し、対象家庭から得ている家庭のおおまかなプロフィール（家族構成と所有機器等）とエキスパートによる判断に基づき行動教師データを生成し、これらに対し本手法を適用し行動推定を行った。その結果、行動把握に重要と考えられる調理と睡眠において多くの家庭で 7 割程度の精度を達成した。

2 行動推定アルゴリズム

主幹電力データは、待機電力を含む各家電の消費電力が一定時間累積した総量の時系列データとして取得される。また、商用の消費電力測定機器の測定間隔は数十秒から数分間隔という低粒度である。したがって微細な消費電力変化やその特徴を捉えることはできず、家電の種別や起動タイミングの推定は困難である。しかし調理、睡眠といった行動においては、主幹電力においても特徴的な電力パターンが出現するため、それを用いて行動内容を推定する。

提案手法では与えられた主幹電力データを一定の時間ウィンドウで分割し、各ウィンドウごとに特徴量を抽出した時系列データを得る。また時間ウィンドウ毎の行動ラベルとして、*out*（外出）、*sleep*（睡眠）、*cook*（調理）および *others*（調理以外の在宅状態）の 4 つのいずれかを与え、これを教師データとして Balanced Random Forest による多クラス分類器を構築する。Balanced Random Forest を用いる理由としては睡眠のように 1 日あたり概ね 8 時間程度を費やす行動と、調理のように 1 日あたり 2 時間程度である行動とではデータ総数に差があるためである。様々な特徴量を試した結果、現ウィンドウ内の電力変化を表すいくつかの指標値、当該ウィンドウの時間帯、およびその直前ウィンドウの行動ラベルを特徴量として用いた場合が最も高精度であったため、本研究ではこれらの特徴量とする。

ここで、家庭毎で睡眠の時間帯や消費電力量は異なるため、複数家庭を統合したデータを用いて分類器を構築すると、異なる行動ラベルに対し似通った特徴量が現れたり、同じ行動ラベルに対し特徴量が大きく異なるといった現象が生じ、学習による分類が困難となる一方、家庭毎の分類器構築は教師データの取得の点で現実的でない。推定対象に合わせた分類器の構築手法として既存の分類器を転用する転移学習 [4] もあるが、いずれにしても対象家庭の教師データは必要とする。行動推定システムを運用する際、行動データの真値を取得することは導入への大きな障壁となる。

これに対し、本稿では、いくつかの家庭の教師データをもとに、家庭毎の分類器を構築し、対象家庭のテストデータからどの分類器が適しているかを選択する手法を提案する。より具体的には、対象家庭のテストデータを選択対象の分類器で分類した結果、ならびにそれら分類器の訓練データを用いる方法を考案している。テストデータに適合しない分類器は睡眠や調理の回数が異常に多いといった、生活として不自然な結果を返すことが多い。また生活パターンが類似する家庭間では、睡眠、外出時間や調理の回数が類似する傾向がある。我々の経験上、分類器に与えるテストデータが適合しない場合、各行動が全行動に占める割合は訓練データのそれと乖離し、適合するのであれば、それらは類似する傾向にある。これを踏まえ、既存のデータで分析を行った結果、行動の存在割合の類似度は精度と

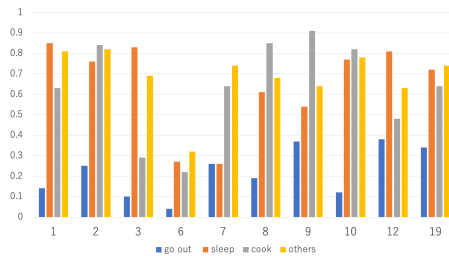


図1 行動ごとの精度の評価

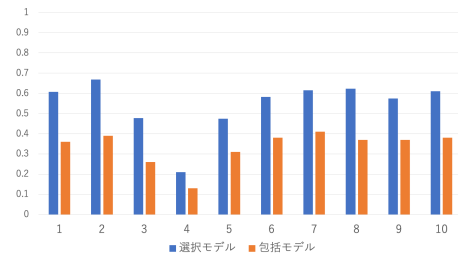


図2 分類器選択の精度の評価

正の相関があることも確認できている。したがって、人間の生活としての自然さと生活パターンの類似度を評価する指標として各行動の存在割合を用いることとした。

ある分類器の訓練データの各行動の行動総数が全行動の行動総数に占める割合を $A_i (i = 0, \dots, n)$ とする。またその分類器を用いてテストデータを判定して得られた推定結果における同様の割合を $B_i (i = 0, \dots, n)$ とする。このとき、対象家庭のテストデータと、用いる分類器の訓練データの類似度を分類器適合度とよび、以下の数式で定義する。

$$Score = \sum_{i=0}^n |1 - \min\{A_i, B_i\} / \max\{A_i, B_i\}|$$

これは、各行動の存在割合の比の1からの乖離度の総和として定義したものである。割合が少ない行動は誤りが精度に及ぼす影響が大きいので、割合の差分の総和ではなく比を用いている。

3 性能評価

提案手法の入力データとなる主幹電力時系列データを取得するため、独居高齢者10家庭の配電盤にクランプを設置し、30秒毎の主幹電力データとリビングおよび寝室に設置した人感センサのデータを191日間取得したデータセットを用いた。行動ラベルに関しては、人感センサの値と我々の知見および対象家庭から得ている家庭のおおまかなプロフィール（年齢性別と所有機器等）に基づき、これら双方のデータを総合して我々自身がラベル付けを実施しこれを真値として用いている。行動ラベルは go out（外出）、sleep（睡眠）、cook（調理）および others（調理以外の在宅状態）としている。

各家庭について、その家庭を除く9家庭分の9個の分類器、およびそれら9家庭全てのデータを学習に用いた分類器の計10個の分類器を作成し、分類器適合度が最も高い分類器を選びそれによる分類結果を行動推定結果とした。データは1家庭あたりおよそ193日分あり（一部欠損している場合もある）、それをテストデータとして利用し検証を実施した。

評価指標として行動毎のf値を用いた結果を図1に示す。

sleep は6家庭で7割以上、others も6家庭でおおよそ7割以上判定できることが確認できた一方、go out の精度は低いことがわかる。これは睡眠時刻は家庭間で大きな差異はないのに対し、職業を持たない高齢者の外出時刻や時間は家庭差が大きいことが原因と考えられる。others についても、電力を使わない行動をとっている場合には go out との判別は難しいため go out と others の誤判定が一定数

見られた。others は絶対数が多いため、go out の多くを others と誤認識しても others の精度は大きく低下することはないものの、go out の精度には大きな影響を与える。

また、分類器選択の効果を確認するため9家庭全てのデータを学習に用いた分類器（包括分類器）と提案手法の分類器選択で選ばれた分類器（最適分類器）の評価を行った。結果を図2に示す。

全ての家庭において、最適分類器は、包括分類器の約1.5～2倍の精度を得ることができ、提案する分類器選択が有用であることが確認できた。全家庭をまとめた包括分類器は、最適分類器以外の分類器と比較しても精度が低く、30秒に1度の主幹電力という低粒度データでは所持家電や生活リズムの異なる複数家庭のデータを与えた場合には適切な分類器を構築できないことも確認できた。またこの分類器選択が最良の分類器（最大f値を達成する分類器）を選んでいるかを4行動のF値の平均で評価したところ、全家庭において、最良分類器と最適分類器が一致したことを確認できた。

4 おわりに

本稿では、低コストかつ非侵襲な行動推定を実現するため、家庭内行動の多くは家電利用が伴うことに着目し、主幹電力を用いた行動推定の手法を提案した。

謝辞 本研究の一部は東北大学電気通信研究所における共同プロジェクト研究の支援によって行われた。

参考文献

- [1] Hoey, J. and Little, J. J.: Value-directed human behavior analysis from video using partially observable Markov decision processes, *IEEE Trans. on pattern analysis and machine intelligence*, Vol. 29, No. 7, pp. 1118–1132 (2007).
- [2] Lee, S.-W. and Mase, K.: Activity and location recognition using wearable sensors, *IEEE pervasive computing*, Vol. 1, No. 3, pp. 24–32 (2002).
- [3] Arakawa, Y., Yasumoto, K., Pattamasirawat, K. and Mizumoto, T.: Improving recognition accuracy for activities of daily living by adding time and area related features, *Proc. of ICMU 2017*, pp. 1–6 (2018).
- [4] Wang, M., Li, W. and Wang, X.: Transferring a generic pedestrian detector towards specific scenes, *Proc. of IEEE CVPR2012*, IEEE, pp. 3274–3281 (2012).