

勾配ブースティング木と並列処理を用いたビル電力需要予測

藤原 佑介[†] 久代 紀之[†] 青山 裕介[†]

九州工業大学[†]

1. はじめに

現在オフィスビルなどでは電力需要の逼迫時のデマンドマネジメントシステムが高まっている。デマンドマネジメントの実現には、電力需要を予測するモデルの構築が必須である。予測モデルからデマンド制御計画の立案が容易であることから、河津らによって異種混合学習技術を用いた線形モデルでのビル電力需要予測ツールの開発^[1] (以下、先行研究) が進められている。一方で、異種混合学習は、予測モデルの構築に時間がかかることによるリアルタイムでの予測モデル構築や、場合分け・説明変数が適切に選択されていない場合に精度が得られないことによる適切な電力需要制御が行えないといった課題がある。

本研究では、事前に場合分け・説明変数の選択が不要でかつ高精度で、並列処理により高速化された勾配ブースティング木を用いたビル電力需要予測方式の開発を行い、リアルタイムでの高精度なモデル構築による電力需要制御の実現をゴールとする。

2. XGBoost

勾配ブースティング木の実装には XGBoost^[2] を使用する。XGBoost は決定木を大量に生成し、各決定木の結果を集計して予測を行う。デマンドマネジメントシステムには、正確な電力需要制御を行うための高精度な電力需要予測と、急な自然災害や気温変化等に対応できる高速なモデル構築が求められる。XGBoost は、予測精度が高く、処理速度が高速で、並列処理を行うことができる。また、モデル構築時に重要となった説明変数を算出でき、予測モデルの精度向上と評価が行いやすいため、正確な電力需要制御が可能となる。

3. 外部条件と電力需要

本研究では、時間や曜日など電力値以外の要素を外部条件と呼ぶ。また、使用する外部条件と予測対象電力は先行研究との比較を行うため、

Building electricity demand prediction using gradient boosting decision tree and parallel processing

[†]Yusuke Fujiwara, [†]Noriyuki Kushi, [†]Yusuke Aoyama

[†]Kyushu Institute of Technology

同様のもの(表 1)を使用する。

また、福田らによる研究^[3]により、全ての外部条件は電力変動に影響を与えていることが確認されている。

表 1 外部条件および予測対象電力

外部条件	<ul style="list-style-type: none"> ・日時(日付、時間、曜日) ・気温 ・推定在室人数 ・継続時間 (始業時から終業時) ・就業日判定 (就業日, 休業日) ・季節区分 (暖房期, 中間期, 冷房期) ・最高気温
予測対象電力	<ul style="list-style-type: none"> ・空調

XGBoost を使用するにあたり、全ての質的変数を量的変数に変換する必要がある。また、量的変数への変換の他に、予測モデルの精度向上のため、日時、就業日判定のデータ成形を行った。

3.1 日時

日時は、日付・時間・曜日の3種類のデータに分けられる。生のデータをそのまま予測モデル構築に使用した場合、質的変数かつ循環性のない外部条件となり、予測精度が悪化する恐れがある。そこで、三角関数の正弦と余弦を用いて日時を使用し、量的変数かつ循環性のある外部条件になるようデータ成形を行い、予測モデルを構築した。

3.2 就業日判定

図 1 に 1 日の各予測対象電力値を示す。今回、電力予測対象として 8 階建てオフィスビルの 5 階を使用する。ビルの電力需要は就業日か休業日かによって大きく異なる。また、用いるデータ期間が 2013 年 4 月 1 日から 2015 年 3 月 31 日の 2 年間 760 日、そのうち就業日が 489 日、休業日が 241 日と約 2 倍の差があり、予測モデルの精度を悪化させる要因となる。今回はデータ数が少ない休業日や不規則な電力値と判断した日を除いた図 1 の右のデータを用いて予測モデル構築と評価を行う。

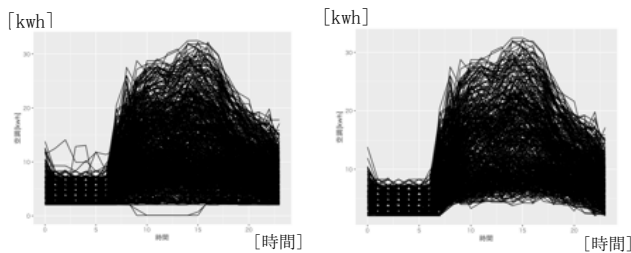


図1 1日の空調電力値
左：2013/04/01～2015/03/31の電力
右：左期間中の就業日のみの電力値

4. 電力需要予測モデル

3.2で記述したデータのうち、2013年4月1日から2014年3月31日(就業日のみ)までのデータを用いて、空調電力値の予測モデルを作成する。

4.1 パラメータチューニング

XGBoostには、モデル構築の際に過学習に関するパラメータが複数存在し、パラメータチューニングにて最適なパラメータ値を決定しなければならない。全ての組み合わせを試す場合、膨大な時間がかかるため、本研究では、ランダムサーチを複数回実行することで、最適なパラメータの算出を行う。また、パラメータの1つである決定木の本数(ブースティング回数)はXGBoostにて最適な値を算出するクロスバリデーション関数があるため、クロスバリデーションを使用して決定木の本数を決定する。

4.2 並列処理

本研究の問題点として、パラメータチューニングに時間がかかることが挙げられる。この処理時間に関する課題の解決方法として、複数PCをネットワーク接続し、複数コアを使用してモデル構築を行うモデル方式を提案する。

5. 評価・考察

3.2で記述したデータのうち、2014年4月1日から2015年3月31日(就業日のみ)までのデータを用いて各測定対象電力値の予測を行った。図2に7月1日と12月1日の消費電力値の実測値と予測値を示す。

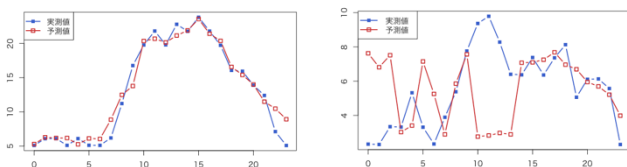


図2 1日の空調電力値予測結果
左：7月1日 右：12月1日
(青：実測値 赤：予測値)

4項で示した予測結果の評価には、先行研究で使われているMAPEを用いる。MAPEの定義式を式(1)に、それぞれのMAPE値と処理時間を表2に示す。

$$MAPE = \frac{\sum |(y_t - \hat{y}_t) / y_t|}{n} \times 100 \quad (y_t \neq 0) \quad (1)$$

y_t : 実測値 \hat{y}_t : 予測値 n : データ数

表2 MAPE値の比較

比較対象	先行研究 (線形モデル)	本研究 (非線形モデル)
空調	22.8	21.1
実行時間	20～30分	約11分

表2から、予測精度は上回ったものの、大きな精度の差は見られなかった。原因として、暖房期・冷房期・中間期といった時期を全て合わせて予測を行なったことが挙げられる。この問題を解決するために、就業日・休業日だけでなく暖房期・冷房期・中間期に分けた予測モデルの構築を行う必要がある。また実行時間は短縮されたものの、より高速化を行わなければゴール達成は難しいと考える。

6. 考察

非線形モデルを使用したにも関わらず、線形モデルでの電力需要精度と差がない要因として、学習データ数とデータの場合分けが考えられる。非線形モデルはモデルの表現力が高い反面、データが多く必要となる。一方で、線形モデルは少ないデータ数で予測を行え、先行研究では異種混合学習によってデータの場合分けを行い、データ数の不足を補っているため、精度に差がないと考えられる。

7. 結論と今後の展開

評価結果より、現状では本研究は線形モデルとの差はあまり見られないと言える。今後も、予測精度の向上と実行時間短縮のため予測モデルの改良と検討を行う予定である。

8. 参考文献

- [1]河津雅忠, 久代紀之 異種混合学習技術を用いたビル電力需要予測ツール 2019
- [2]Tianqi Chen, Carlos Guestrin XGBoost: A Scalable Tree Boosting System 2016
- [3]福田亜美, 妻鹿利宏, 久代紀之 BEMSデータを用いたビル内電力需給予測モデル 2016