An LSTM-Based Online Prediction Method for Building Electric Load During COVID-19

Hao Tu, Srdjan Lukic
FREEDM Systems Center, North Carolina State University
htu@ncsu.edu  smlukic@ncsu.edu

Abhishek Dubey, Gabor Karsai
Institute for Software Integrated Systems, Vanderbilt University
abhishekdubey@vanderbilt.edu  gabor.karsai@vanderbilt.edu
Motivation

- Accurate prediction of building electrics load is important to its efficient operation and necessary for some advanced functionalities
  - Frequency regulation
  - Virtual battery by exploiting the thermal capacity

- Challenges
  - Detailed building model and parameters are difficult to obtain
  - Load pattern can change during events like COVID-19

- Proposed method: LSTM-based online learning
  - Machine learning method only requires history load data
  - Online learning method can adapt to load pattern change automatically
Given a variable of interest $y$ and a feature vector $x$, the single-step time series prediction problem is to learn a nonlinear mapping function $F$ that uses the historical sequence $\{x_k, x_{k-1}, \ldots, x_{k-l+1}, x_{k-l}\}$ as input to make prediction for the next time step $\tilde{y}_{k+1}$,

$$\tilde{y}_{k+1} = F(x_k, x_{k-1}, \ldots, x_{k-l+1}, x_{k-l})$$

Where $y_{k+1}$ is the variable of interest at time step $k + 1$ and $\tilde{y}_{k+1}$ is its prediction. $x_k$ is a vector that contains the input information at time step $k$. Note $x_k$ commonly contains the true value of $y_k$ and some other selected features.
Long short-term memory (LSTM) is a type of recurrent neural networks that is widely used for time series prediction. Its unique structure featuring input, forget, and output gates helps to pass information between steps.

\[
\begin{align*}
  i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\
  f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\
  o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\
  \tilde{C}_t &= \tanh(W_C x_t + U_C h_{t-1} + b_C) \\
  C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\
  h_t &= o_t * \tanh(C_t)
\end{align*}
\]
Identified features for LSTM model

- Historical electric load data $y_k$
- Outside air temperature $T_{OAT}$
- Hour of the day $H$
  - $H \in \{0,1\}^{24}$ is a 24-dimensional one-hot encoded vector
- Day of the week $D$
  - $D \in \{0,1\}^7$ is a 7-dimensional one-hot encoded vector
- Day type $G$
  - $G \in \{0,1\}^1$ is binary feature, $G = 1$ if it is a holiday and $G = 0$ otherwise

The input feature is the concatenation of above features,

$$x_k = \text{concat}(y_k, T_{OAT}, H, D, G)$$
Online Learning steps

1. Make prediction
   \[ \tilde{y}_{k+1} = F_k(x_k, x_{k-1}, \ldots, x_{k-l+1}, x_{k-l}) \]

2. Receive new data ground truth \( y_{k+1} \)

3. Retrain model with gradient descent method,
   \[ W_{k+1} = W_k - \alpha \nabla \text{loss}(W_k, y_{k+1}, \tilde{y}_{k+1}) \]
   Where \( W_k \) is the set of all the parameters of model \( F_k \) at time step \( k \).

Online learning has the capability of adapting the model to new incoming data and discovering the underlying new pattern if a concept change has happened. However, it is difficult to predetermine the optimal value for learning rate \( \alpha \).
Online Learning with Adaptive Learning Rate

- **Slow learner** $\alpha_1$
  - Predict $\hat{y}_{k+1}^1$ using $F_k$ and $\{x_k, ..., x_{k-l}\}$

- **Average learner** $\alpha_2$
  - Predict $\hat{y}_{k+1}^2$ using $F_k$ and $\{x_k, ..., x_{k-l}\}$

- **Fast learner** $\alpha_3$
  - Predict $\hat{y}_{k+1}^3$ using $F_k$ and $\{x_k, ..., x_{k-l}\}$

Receive ground truth $y_{k+1}$

Compare $y_{k+1}$ and $\hat{y}_{k+1}^1$, $\hat{y}_{k+1}^2$, $\hat{y}_{k+1}^3$ and replicate most accurate model three times

Increase or decrease the learning rates by $\delta \alpha$ or not change

Train model with $y_{k+1}$, $\{x_{k+1}, ..., x_{k-l+1}\}$ and $\alpha_1$

Train model with $y_{k+1}$, $\{x_{k+1}, ..., x_{k-l+1}\}$ and $\alpha_2$

Train model with $y_{k+1}$, $\{x_{k+1}, ..., x_{k-l+1}\}$ and $\alpha_3$

$k = k + 1$

$\alpha_1 < \alpha_2 < \alpha_3$
Algorithm 1 Online LSTM with Adaptive Learning Rate

1: procedure PredictAndUpdate
2:   for $i = 1 \rightarrow 3$ do
3:     $\{\tilde{y}_{k+1}^{(i)}\} \leftarrow \text{PredictLearner}_i(x_k, ... x_{k-1}, W_k^{(i)})$
4:     $\tilde{y}_{k+1} = (\tilde{y}_{k+1}^{(1)} + \tilde{y}_{k+1}^{(2)} + \tilde{y}_{k+1}^{(3)}) / 3$
5:   
6:   while ($y_{k+1}$ is not available) do
7:     Wait
8:   
9:   for $i = 1 \rightarrow 3$ do
10:      $err^{(i)} = |\tilde{y}_{k+1}^{(i)} - y_{k+1}|$
11:   
12:   $\text{IdxBestLearner} \leftarrow \text{index for lowest error } err^{(i)}$
13: 
14:   if $\text{IdxBestLearner} == 1$ then
15:     $\{\alpha_1, \alpha_2, \alpha_3\} \leftarrow \{\alpha_1 - \delta\alpha, \alpha_2 - \delta\alpha, \alpha_3 - \delta\alpha\}$
16:     $\text{BestLearner} \leftarrow \text{SlowLearner}$
17:   else if $\text{IdxBestLearner} == 2$ then
18:     $\{\alpha_1, \alpha_2, \alpha_3\} \leftarrow \{\alpha_1, \alpha_2, \alpha_3\}$
19:     $\text{BestLearner} \leftarrow \text{AverageLearner}$
20:   else if $\text{IdxBestLearner} == 3$ then
21:     $\{\alpha_1, \alpha_2, \alpha_3\} \leftarrow \{\alpha_1 + \delta\alpha, \alpha_2 + \delta\alpha, \alpha_3 + \delta\alpha\}$
22:     $\text{BestLearner} \leftarrow \text{FastLearner}$
23:     $\{\text{SlowLearner, AverageLearner, FastLearner}\} \leftarrow \{\text{BestLearner, BestLearner, BestLearner}\}$
24: 
25:   for $i = 1 \rightarrow 3$ do
26:     TrainLearner$_i$ on $(x_k, ... x_{k-1}, y_{k+1})$ with $\alpha_i$
Data Set Analysis

Electric load data for a building in California from July 2019 to July 2020
Baselines

Two basslines:
1. Off-line LSTM (Off-LSTM): this is LSTM model without online training. In this case, the data from 6.30.2019 to 3.5.2020 are used for training.
2. Online LSTM with fixed learning rate (On-LSTM-FLR): takes the trained off-line LSTM model and update the model on new samples every time step with a fixed learning rate $\alpha = 0.02$.

Proposed method: takes the trained off-line LSTM model and update the model as described.
Experiment results:

<table>
<thead>
<tr>
<th>Phase Week</th>
<th>Pre-COVID-19</th>
<th>COVID-19</th>
<th>Post-COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.9-3.15</td>
<td>3.16-3.22</td>
<td>3.23-3.29</td>
</tr>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>35.9</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>31.3</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>30.5</td>
</tr>
</tbody>
</table>

Performance metric: mean absolute error (MAE):

\[
MAE = \frac{1}{N} \sum_{k=N_0+1}^{N_0+N} |\tilde{y}_k - y_k|
\]

Where \( N \) is the number of predictions and is selected to be 168, which corresponds to one-week long hourly data.
Experiment results:

<table>
<thead>
<tr>
<th>Phase</th>
<th>Pre-COVID-19</th>
<th>COVID-19</th>
<th>Post-COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week</td>
<td>3.9-3.15</td>
<td>3.16-3.22</td>
<td>3.23-3.29</td>
</tr>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>35.9</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>31.3</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>30.5</td>
</tr>
</tbody>
</table>

Results for pre-COVID-19 phase, week from Mar 9 2020 to Mar 15 2020
Experiment results:

<table>
<thead>
<tr>
<th>Phase</th>
<th>Pre-COVID-19</th>
<th>COVID-19</th>
<th>Post-COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.9-3.15</td>
<td>3.16-3.22</td>
<td>5.25-5.31</td>
</tr>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>76.2</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>69.1</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>67.5</td>
</tr>
</tbody>
</table>

Results for COVID-19 phase, week from Mar 16 2020 to Mar 22 2020
## Results

### Experiment results:

<table>
<thead>
<tr>
<th>Phase</th>
<th>Pre-COVID-19</th>
<th>COVID-19</th>
<th>Post-COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.9-3.15</td>
<td>3.16-3.22</td>
<td>5.11-5.17</td>
</tr>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>41.3</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>34.9</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>32.4</td>
</tr>
</tbody>
</table>

Results for COVID-19 phase, week from May 11 2020 to May 17 2020

- (a) Off-LSTM, MAE = 41.3 kW
- (b) On-LSTM-FLR, MAE = 34.9 kW
- (c) Proposed method, MAE = 32.4 kW
### Results

**Experiment results:**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>35.9</td>
<td>43.0</td>
<td>41.3</td>
<td>54.8</td>
<td>76.2</td>
<td>62.0</td>
<td>59.1</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>31.3</td>
<td>33.4</td>
<td>34.9</td>
<td>51.3</td>
<td>69.1</td>
<td>59.7</td>
<td>53.0</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>30.5</td>
<td>31.4</td>
<td>32.4</td>
<td>44.4</td>
<td>67.5</td>
<td>55.9</td>
<td>51.5</td>
</tr>
</tbody>
</table>

Results for post-COVID-19 phase, week from May 25 2020 to May 31 2020
Experiment results:

<table>
<thead>
<tr>
<th>Phase Week</th>
<th>Pre-COVID-19 3.9-3.15</th>
<th>3.16-3.22</th>
<th>3.23-3.29</th>
<th>COVID-19 5.4-5.10</th>
<th>5.11-5.17</th>
<th>5.18-5.24</th>
<th>Post-COVID-19 5.25-5.31</th>
<th>6.1-6.7</th>
<th>6.8-6.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-LSTM</td>
<td>31.5</td>
<td>34.3</td>
<td>35.9</td>
<td>43.0</td>
<td>41.3</td>
<td>54.8</td>
<td>76.2</td>
<td>62.0</td>
<td>59.1</td>
</tr>
<tr>
<td>On-LSTM-FLR</td>
<td>35.0</td>
<td>35.5</td>
<td>31.3</td>
<td>33.4</td>
<td>34.9</td>
<td>51.3</td>
<td>69.1</td>
<td>59.7</td>
<td>53.0</td>
</tr>
<tr>
<td>Proposed</td>
<td>31.7</td>
<td>30.9</td>
<td>30.5</td>
<td>31.4</td>
<td>32.4</td>
<td>44.4</td>
<td>67.5</td>
<td>55.9</td>
<td>51.5</td>
</tr>
</tbody>
</table>

Conclusion:
The proposed method can quickly adapt the model to the concept changes during COVID-19 and reduce the prediction errors.
Thank you for your time.