Transformers and Contextual Information in Temperature Prediction of Residential Buildings for Improved Energy Consumption

Sai Sushanth Varma Kalidindi, Hadi Banaee, Amy Loutfi

Center for Applied Autonomous Sensor Systems (AASS)
Örebro University, Örebro, Sweden
{sushanth.kalidindi, hadi.banaee, amy.loutfi}@oru.se

Abstract

Energy optimization plays a vital role in decreasing the carbon footprint of residential buildings. In this paper, we present a prediction model of indoor temperature in residential buildings in three different case studies in different towns in Sweden. To predict the indoor temperature accurately, a dataset based on several years of data collection (up to 7 years) has been used. This paper applies both the traditional LSTM model as well as the more recent transformer model. The latter has been used because of its ability to perform a mechanism of self-attention that shows particular promise in multivariate sensor data. In addition to these algorithms, the data set is also modified based on contextual information and compared against an approach where no contextual information is used. Contextual information in this case takes into account the physical location of specific apartment units within the full residence and builds individual models based on the location of the unit. The results demonstrate that transformers are better suited for task of prediction, and that transformers combined with contextual information, provide a suitable approach for energy consumption prediction.

Introduction

Residential buildings still use more than half of the energy consumption in the Nordic region in heating, ventilation, air conditioning and electricity (Energimyndigheten 2021). Most of the energy produced for heating is generated by fossil fuels and biomass (Jenni Patronen and Torvestad 2017). Therefore, it is important to optimize the energy consumption. For developing such an efficient strategy, prediction of indoor temperature plays an important role and can save up to 5 - 15% of energy in existing buildings (Afroz et al. 2018). This paper deals with multivariate time series since it involves several variables that impact the prediction of *indoor temperature* over a period of time. Several AI methods can be used for time series prediction, but there is a need to investigate new methods that can account for the contextual information of building and long data sequences. In this paper, we address the feasibility of predicting indoor temperature time series data with transformers and compared to LSTM as a benchmark. And, we investigate the role of contextual information such as *size*, *location of the apartment*, *weather*, *seasons*, *occupancy*, and *passive heating or cooling* for prediction of indoor temperature.

Background

Various AI techniques have been used in the literature to predict the indoor temperature in buildings. The analytical-based models are poor in prediction due to many assumptions and parameters (Braun and Chaturvedi 2002). On the contrary, data-driven models which completely rely on experimental data provide appropriate prediction accuracy if enough data is available (Afram and Janabi-Sharifi 2015).

For the prediction of indoor temperature time series data, it is shown that the prediction models based on data-driven modeling techniques, i.e., Neural networks, are more effective than analytical models (Soleimani-Mohseni, Thomas, and Fahlen 2006; Ruano et al. 2006).

However, a few works have considered modelling indoor temperature time series prediction of up to 28 Hrs and energy consumption prediction of up to three months (Ruano et al. 2006; Somu, MR, and Ramamritham 2020; Soleimani-Mohseni, Thomas, and Fahlen 2006; Hietaharju, Ruusunen, and Leiviskä 2018). In general, despite the efforts in *Artificial Neural Networks* (ANN) and *Recurrent Neural Networks* (RNN), these models still lack the predictions for long sequences of data and complex features (Jain et al. 2014) due to vanishing gradients and low memory.

Moreover, a new approach based on attention mechanisms called *transformer* has been introduced to model the sequences in Natural Language Processing (NLP) (Vaswani et al. 2017). transformer-based approaches work well on processing long data sequences and complex patterns (Wu et al. 2020), which make it a proper option for indoor temperature prediction. There is not much research are done using transformers to predict particularly *indoor temperature*. One of the works which use transformers to predict energy consumption is (Rao and Zhang 2020) which predicts one step ahead. Besides the time series data of indoor temperature, contextual information provides a broader understanding of the system and a better prediction accuracy in the model. An example is to include the season context (Afroz et al. 2018), wherein the data was split accordingly to prove that including context provides a better prediction accuracy in long-term forecasting. In this paper, we aim to address the prediction of data using LSTM and transformer techniques, as well as considering the significance of the context of *floors* of the buildings in the accuracy of the prediction.

Method

The method we propose is to 1) apply both LSTM and transformer models each individually for predicting indoor temperature time series in a building, and 2) include the contextual information to present a context-aware transformer model for prediction. The overall comparison of the proposed methods is processed by single-step and multi-step prediction of indoor temperature by LSTM and transformer as shown in Figure 1. Two approaches are followed to model the methods, one is training the prediction models for each building with the complete climatic sensor data (*overall* model) and the other is to include the context of floors to model a transformer based approach (*context-aware* model).

Input Information

The data was measured in residential apartment buildings across Sweden provided by EcoGuard AB. An illustration of a sample building is shown in Figure 2 with all the apartments as the individual units. For the privacy of the residents, the buildings used in this research are not shown.

Indoor climate sensors were placed on the walls in the living area at a height of 1.7 m above the floor. The temperature sensor measurements have been taken every 15 minutes for 2-7 years based on the building and availability. The temperature outside the building has been also measured for the same period at every 15 minutes.

Three buildings have been investigated in this paper, where their information is presented in Table 1. For all the buildings, the heat is provided through a district heating network. The control system in a building operates the heating temperature based on the average indoor temperature in the building.

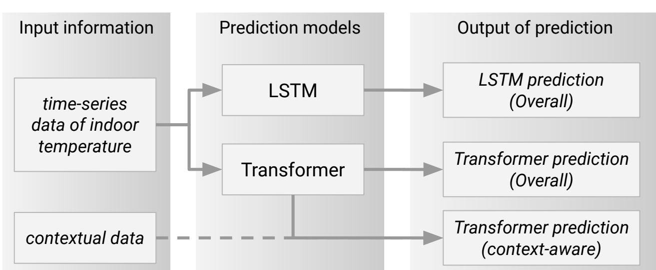

Figure 1: Representation of modeling methods in buildings including context.

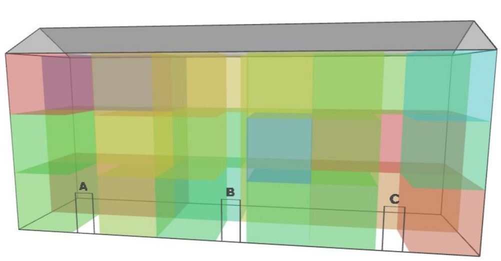

Figure 2: An illustration of a building (27 units in 3 floors).

Residential buildings in Sweden			
Name of city (Location of building)	Number of floors	Number of apartments	Measurement Period
Borås	5	21	2014 Aug – 2021 Oct (7y,2m)
Örebro	9	141	2017 Oct – 2021 July (3y, 9m)
Piteå	3	15	2019 June – 2021 Oct (2y,4m)

Table 1: Input information of the buildings to be modelled.

In the first approach, the input of the prediction models is alike with 250k samples in Borås, 131k samples in Örebro, and 79k samples in Piteå. In the second approach, with the context of floors, the input is split into three layouts for each building: the apartments on the ground floor to *ground-layout*, the apartments connected to the roof to *top-layout*, and the rest of the floors to *middle-layout*. Individual models are trained for each layout for each of the buildings. Contextual information in this case takes into account the physical location of specific apartment units within the full residence and builds individual models based on the location of the unit. The *ground-layout* and *top-layout* split individually due to variations of temperature from ground and the roof of building.

Prediction Models

LSTM In general, LSTM employs three gates: input gate, forget gate, and output gate to modulate the information across the cells. We apply the common architecture of LSTM (Hochreiter 1997), but we tune the overall model for our specific case study, as follows. The parameter setting of the LSTM is similar for all three buildings: two LSTM layers with 50 neurons each, and a Relu activation function. We use Dropout of 0.2 for each layer, batch size of 150, a learning rate of 0.001, optimizer Adam, and loss function mean absolute error (MAE). All three models are trained for 40 epochs. Also, look back of 128-time steps is used for predicting the future time steps of indoor Temperature.

The next future indoor temperature is predicted for 15 min, 24 hrs, and 72 hrs.

Transformer The proposed transformer model follows the original transformer architecture (Vaswani et al. 2017) consisting of encoder and decoder layers. The encoder is composed of a stack of N=5 identical layers. Each layer has two sub-layers: a multi-head self-attention mechanism, and a positional wise fully connected feed-forward network. Positional wise encoding with sine and cosine functions is used to encode sequential information in time series data by element-wise addition of input vector with a positional encoding vector. Each sub-layer is followed by a normalization layer. The encoder produces a dimensional vector, d_{model}, to feed to the decoder (Vaswani et al. 2017; Wu et al. 2020).

The decoder is also composed of a stack of N=5 identical layers. In addition to the two sub-layers, the decoder inserts a third sub-layer to apply multi-head attention on the output of the encoder stack. The output layer maps the output of the last decoder to the target time sequence. The output embeddings are offset by one position to ensure that the prediction only depends on the past data points (Vaswani et al. 2017; Wu et al. 2020). For our case study, the parameter setting of the transformer model is similar to the LSTM model, but without a dropout in the encoder and decoder.

Evaluation In order to evaluate the prediction accuracy of the models, root mean square error (RMSE) is used as the evaluation criteria of the models The RMSE calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{1}$$

where $\hat{y}_i$ is the predicted value and y_i is the true value, and N is the sample size. The smaller the value of RMSE, the better the forecasting.

Results

The results are presented in two phases, 1) the overall modeling LSTM and transformer for each of the buildings, and 2) modeling a transformer for each of the contexts of *floor*. The test data for all the models are from November 2020 to June 2021 to have a similar test set and an equal comparison.

Phase 1: Each building in Borås, Örebro, and Piteå has been modelled by LSTM and transformer. The prediction results of these models are presented in Table 2. The models are evaluated with different prediction lengths: 1 step (15min), 96 steps (24hrs), and 288 steps (72 hrs), and their performance was evaluated by calculating the average RMSE of each apartment. A sample of results prediction models are presented in Figure 3 for 72 hours (288 time steps) ahead. The apartment is randomly selected from the 4th floor of the building in Örebro. The top graph shows transformer performs better than LSTM in this example. The bottom one shows the significance of involving the context of floor in transformer model by comparing the results of the overall versus the context-aware transformer models.

The average RSME results of LSTM and transformer for all the three buildings are shown in Table 2. Considering these results, transformer outperforms LSTM in all three prediction scenarios (15 min, 24 Hrs, 72 Hrs). In average, the performance is increased 80% in 1 time step, 75% in 96 time steps, and 34% in 288 time steps.

Phase 2: This phase has trained both LSTM and transformer model for contexts of floors in Borås and Örebro. Due to the lack of data from the floors in Piteå, the model is

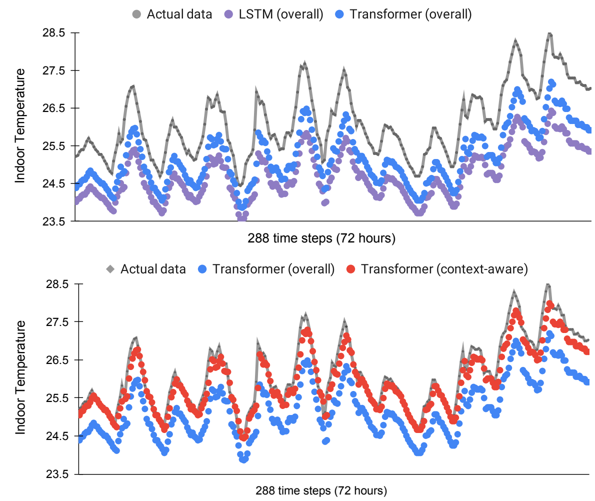

Figure 3: Indoor temperature prediction in a 4th floor unit in Örebro for 72 hours: overall LSTM vs overall transformer (top). overall vs. context-aware transformer (bottom).

Prediction		RMSE (average)		
Building location	Model	1 step (15 min)	96 steps (24 Hrs)	288 steps (72 Hrs)
Borås	LSTM	0.049	0.078	0.121
	Transformer	**0.009**	**0.023**	**0.093**
Örebro	LSTM	0.046	0.082	0.175
	Transformer	**0.007**	**0.017**	**0.091**
Piteå	LSTM	0.066	0.126	0.237
	Transformer	**0.012**	**0.049**	**0.158**

Table 2: Comparison of the average RMSE using LSTM and transformer models in different buildings.

not presented. The results are presented only for the transformer models because the improvement is similar over LSTM in this phase and due to the page limitation. Figure 4 presents the results for the second approach by training the model for different layouts in the locations of Borås and Örebro. The models are trained by LSTM and transformer but the results presented in the Figure 4 are for the transformer. In both locations, considering the context of floors and modeling according to the layout, the models have greater accuracy. Figure 4 also compares the context result with the overall model as the first approach (grey bar).

Implementation details: All the experiments are carried out under the running environment of Intel i7-10870H 2.2-4.8GHz, 32GBs of RAM, NVIDIA GeForce RTX 3080 16Gbs, and Windows 10.

Conclusions

In this short paper, we presented a transformer based approach to predict time series data on buildings. Compared to LSTM, this approach leverages the self-attention mechanisms to model the sequence data and learns the complex dependencies from time series data. Our results suggest that

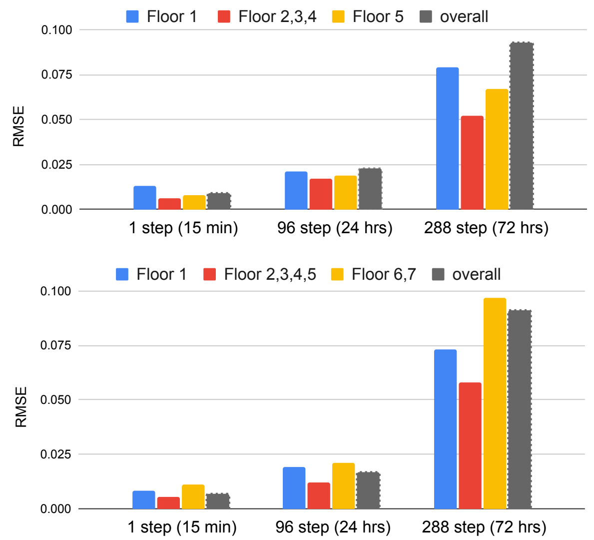

Figure 4: The average RMSE using transformer for overall and context of floors in Borås (top), and Örebro (bottom).

including the contextual information (floors in our case) improves the model performance. However, the improvement when considering the context of floors is significant but considering other contextual information in the training step is likely to improve the model much further.

Although this paper focuses on modeling overall building and in the context of floors, we hypothesize that our approach can be further extended to include the *location*, *weather*, *seasons*, *occupancy*, and *passive heating or cooling* in the training. The future work of this research can be a better formalization and integration of context in the learning process, to better predict the long-term results. Moreover, we will address the optimization problem of the building using the proposed prediction models.

Acknowledgments

This work has been supported by the Industrial Graduate School Collaborative AI & Robotics funded by the Swedish Knowledge Foundation Dnr:20190128.

References

Energimyndigheten. 2021. Energy in Sweden. *Energy*, (ET 2021: 11): 1–18

Hietaharju, P.; and Ruusunen, M. 2015. A concept for cutting peak loads in district heating. *Proceedings of the Automaatio XXI, Helsinki, Finland*, 17–18

Hietaharju, P.; and Ruusunen, M. 2018. Peak load cutting in district heating network. In *Proceedings of EUROSIM2016/SIMS2016*, Linköping Electronic Conference Proceedings, 99–104. Linköping University Electronic Press

Jenni Patronen, E. K.; and Torvestad, C., eds. 2017. *Nordic heating and cooling: Nordic approach to EU's Heating and Cooling Strategy*. Copenhagen: Nordisk Ministerråd

Ståhl, F. 2009. *Influence of thermal mass on the heating and cooling demands of a building unit*. Chalmers Tekniska Hogskola (Sweden)

Afroz, Z.; Urmee, T.; Shafiullah, G.; and Higgins, G. 2018. Real-time prediction model for indoor temperature in a commercial building. *Applied energy*, 231: 29–53

Braun, J. E.; and Chaturvedi, N. 2002. An inverse gray-box model for transient building load prediction. *HVAC&R Research*, 8(1): 73–99

Afram, A.; and Janabi-Sharifi, F. 2015. Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods. *Energy and Buildings*, 94: 121–149

Soleimani-Mohseni, M.; Thomas, B.; and Fahlen, P. 2006. Estimation of operative temperature in buildings using artificial neural networks. *Energy and Buildings*, 38(6): 635–640

Ruano, A. E.; Crispim, E. M.; Conceiçao, E. Z.; and Lúcio, M. M. J. 2006. Prediction of building's temperature using neural networks models. *Energy and Buildings*, 38(6): 682–694

Jain, R. K.; Smith, K. M.; Culligan, P. J.; and Taylor, J. E. 2014. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Applied Energy*, 123: 168–178

Somu, N.; MR, G. R.; and Ramamritham, K. 2020. A hybrid model for building energy consumption forecasting using long short term memory networks. *Applied Energy*, 261: 114131

Hochreiter, S. 1997. JA1 4 rgen Schmidhuber (1997)."Long Short-Term Memory". *Neural Computation*, 9(8)

Wu, N.; Green, B.; Ben, X.; and O'Banion, S. 2020. Deep transformer models for time series forecasting: The influenza prevalence case. *arXiv preprint arXiv:2001.08317*

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. In *Advances in neural information processing systems*, 5998–6008

Adebiyi, A. A.; Adewumi, A. O.; and Ayo, C. K. 2014. Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014

Afroz, Z.; Urmee, T.; Shafiullah, G.; and Higgins, G. 2018. Real-time prediction model for indoor temperature in a commercial building. *Applied energy*, 231: 29–53

Rao, Z.; and Zhang, Y. 2020. Transformer-based power system energy prediction model. In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, 913–917. IEEE