Smart Hot Water Control with Learned Human Behavior for Minimal Energy Consumption

Tim Sonnekalb[1,2] and Sergio Lucia[1]

[1]Technische Universität Berlin, Laboratory of Internet of Things for Smart Buildings,
Einstein Center Digital Future, 10587 Berlin, Germany
Email: sergio.lucia@tu-berlin.de

[2]German Aerospace Center (DLR), Institute for Data Science, 07745 Jena, Germany
Email: tim.sonnekalb@dlr.de

***Abstract*—This work presents an approach to automatically adapt domestic hot water heaters to to individual human behavior based on real IoT data. For this purpose, a large collection of data from domestic hot water heaters is analyzed to learn the consumption behaviors of each user. The human behavior is learned using two different approaches that we compare: neural networks and Gaussian processes with periodic kernels.**

The learned human behavior is used to create an optimal hot water schedule that adapts to each user and thus saves between 20 and 34% of the energy used with a default schedule. We also propose an eco-parameter so that each user can determine a trade-off between maximum comfort (always having hot water available) and maximum energy savings.

I. Introduction

Optimizing the energy used for hot water heating (not including space heating) is a relevant problem because the domestic hot water (DHW) production amounts for approximately 17% of the total energy consumption in the residential sector in the USA and 14% in the European Union [1]. The total energy consumption of the residential sector is in turn very dependent on the country, which is mostly attributed to the prevailing climate in the region and it raises yearly with increasing population. For example, it represents the 16% of the total energy consumption in Finland (the minimum) over 25% in the USA to 50% of the total energy consumption in Saudi Arabia [2]. In the EU, the average energy consumption in the residential sector amounts to around 25% [3].

Current models of DHW heaters keep hot water available in fixed time periods. A typical hot water heater, as it will be considered in this paper, has a preset time window from 05:30 AM to 10:00 PM every day in delivery condition. In this time window from morning to night, it is ensured that hot water is always available but there is rarely a user who needs hot water the whole day. To allow control possibilities of the heating schedule, the manufacturer of the heaters considered in this work enabled the definition of different time periods that can be entered for different weekdays. However, in practice users often do not take advantage of this possibility and use the standard configuration after the installation of their water heater. Since only a small proportion of users actually adapts the time window to their needs, there is a large potential for energy savings if such an adaptation could be performed automatically based on IoT data. Achieving such energy savings is the main motivation of this work.

978-1-5386-4980-0/19/$31.00 ©2019 IEEE

Learning human behaviour based on data has been recently studied in several works in the context of smart homes. The heating times of a solar-combined system of one household were predicted using artificial neural networks (ANNs) [4]. The authors of the work use took into consideration the ambient temperature, the input and output temperature of the DHW production system and the volume flow. In this work, we want to limit the data to only device-internal variables. Other approaches predict only the future measurements of sensors for heating, lighting, ventilation or water heating to estimate the house inhabitants behavior and derive their time schedules [5].

Approaches in [6] and [7] uses also neural networks for domestic hot water energy prediction. However, the authors use data sources which are outdated due to the further development of the heating devices. Gaussian Processes as machine learning algorithm are used to predict the external temperature with the goal of minimizing the heating system costs in [8], but only weather and electricity grid data is used, disregarding an adaptation to individual users. Gaussian Processes are a suitable tool predict temperature profiles as was also shown in [9] for creating a thermal building model.

Behavioural patterns and user profiles on personal preferences and demographic facts are investigated statistically for space heating in [10], which can be also used to formulate energy-saving policies. In [11], weather and building data were used to analyze the influence on energy savings but the main challenge is the availability of accurate data.

The main contribution of this work is the use and evaluation of real world data sets that depend only on device-internal measurements to learn human behavior using both Gaussian processes and neural networks. Another contribution of this work is to propose an approach to generate optimal heating schedules based on learned human behavior that can obtain an optimal trade-off between human comfort and energy consumption based on the user's needs. Apart from achieving significant energy savings, we believe that the proposed methods can be integrated in other smart energy systems that make use of models to perform optimal decisions [12].

II. Case Study: Hot Water Control

We describe in this section the data and scenarios considered for the development of a smart hot water control system. Smart in this context means that the control unit of the hot water heating system adapts the heating times to the behavior of the user and it is therefore neither static nor manually configured.

A. Real World Data sets

For this work, an European heating system company has provided 17 different data sets that will be used to develop and evaluate the proposed approach. The data sets come from 17 different and independent DHW devices of the same type. The data is available in a time period of one and a half years from October 2016 to March 2018. The devices are powered by gas with a capacity of 19 kW and their integrated water reservoir holds an amount of 100 liters. The data sets were recorded following an event-based scheme via the smart-home gateway in the devices, meaning that every time a sensor measures a new value, a message is sent from the gateway to the manufacturer's back-end. These events are converted to a data set with data examples at a distance of one minute. The data sets used in this application include only device-internal measurements from internal sensors of the device and an additional external outside temperature sensor. The features of the data set are the measurement of temperature in the water reservoir and from the outside, time (extracted to minutes, weekday and season) and the power rate of the boiler in percentage over the maximum possible power. Additionally, the data contains device-internal states as Boolean expression of the storage charge pump which is used in the pre-processing sec. III-B to delimit hot water heating from space heating of the combined devices.

B. Typical User Behavior

Making predictions about the future user behavior requires that the times in which the user needs warm water occur in regular patterns. The data shown in Fig. 1 on the real data exemplary for one user shows the potential for energy savings if a personalized automatic heating schedule can be obtained. The power rate at which the boiler operates, called *modulation* (black curve) is directly linked to the consumed energy (see Eq. 3). The negative temperature gradient (blue bars) is used to detect water extractions out of the boiler because the available data is only device-internal and it does not include any flow measurements. Both quantities are averaged over one year.

The most of the hot water is consumed from 4 PM to 8 PM, while during the previous hours from 7 AM to 3 PM a consistent low temperature gradient indicates that no or very small amounts of water are needed in this period. The constant high values of the power consumption curve from 5:30 AM to 10:00 PM are due to the standard heating times set by the manufacturer and were obviously not changed by the user over the whole year.

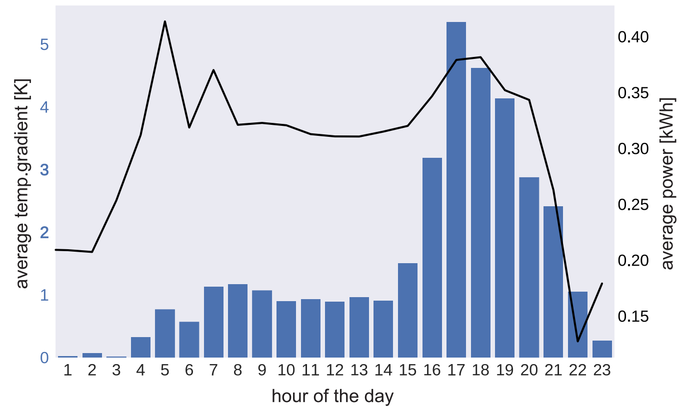

Fig. 1. Negative temperature gradient (blue bars) and power consumption (black line) of the water heater for one day (average over one year).

III. Learned Human Behavior

A. Structure of the proposed approach

The main contribution of this work is to learn the human behavior related to hot water needs so that the heating times can be automatically adapted, leading to significant energy savings. To do so, we developed a data analysis process that is shown in Fig. 2. It includes two machine learning-based models obtained from real-data of 17 different heaters. The blue-colored inputs are the domestic hot water heater data from the respective heating device / household, split into two sets as inputs for the model.

The temperature trend of the water reservoir is the data variable used to separate the user's behavior and the behavior

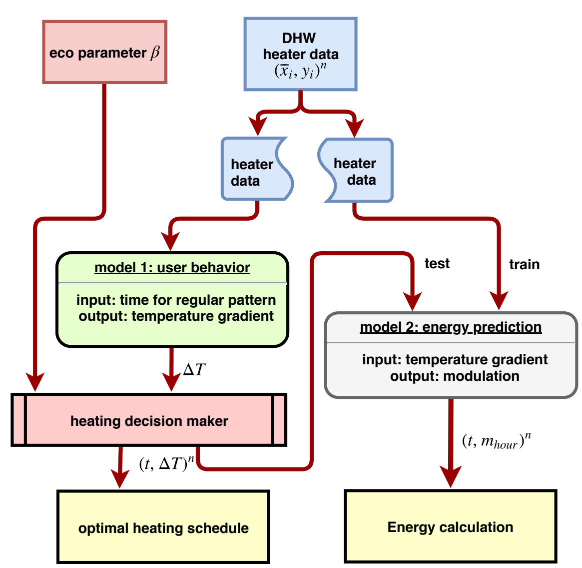

Fig. 2. Data analysis process for smart hot water control with learned human behavior.

of the heater. It was established that all positive temperature trends are caused by the heating behavior and all negative temperature trends, except the small ones, are representing a water extraction and thus they correspond to the user's behavior. The small negative temperature gradients occur due to the unavoidable cooling of the water despite isolation of the boiler. The threshold between ambient cooling and water extraction was defined empirically. Under the assumption that cooling occurs much more often than water extractions the threshold was determined via a histogram to a temperature gradient, for example $\Theta_c = 0.04$ K/s for one of the data set. This threshold is individual per heater because of its installation location within the house, it should be larger for the devices position at cooler places. No meta information for the installation place is given within the IoT data.

The model 1 learns the user behavior with different time measures as input (the minutes of the day, the weekday and the season) with the negative temperature gradient as measure for the amount of extracted water as target. Model 2 is mainly built for evaluation of this work and not necessary for the use in practice and is an emulation of the effects that the computed optimal schedules would have when applied to the domestic heaters. Future work includes the evaluation on real devices. Model 2 performs an energy prediction, it needs the positive temperature gradient as input and predicts the modulation which represents how much energy is needed to achieve certain temperature increase in the water reservoir.

Mathematically, both models can be formulated as:

$$\text{model 1: } f_{\text{model1}}(t_{\text{min}}, t_{\text{we}}, t_{\text{sea}}, T_{\text{out}}, T_{dem}^{-i}) = -(\Delta T_{\text{dem}} - \Theta_c), \tag{1}$$

$$\text{model 2: } f_{\text{model2}}(\Delta T_{\text{dem}}, \Delta T_{\text{dem}}^{-i}, m_{\text{hour}}^{-i}) = m_{\text{hour}}. \tag{2}$$

Both functions f_{model1} and f_{model2} represent the black-box functions of the machine learning functions with time inputs minutes of the day t_{min}, weekday t_{we}, the season t_{sea}, the outside temperature T_{out} and the recurrent temperature gradients or recurrent modulations T_{dem}^{-i}, m_{hour^i} with $m \in \{1, 2, 3\}$, the modulation m_{hour}, a threshold Θ_c for physical cooling and temperature gradient ΔT_{dem} as output. The target values include a negative sign to predict positive values. The outputs of model 1 are used in the heating decision maker to derive an optimal heating schedule. Using an eco-parameter β, the user can set a preference towards maximum energy saving or maximum comfort. The parameter acts as a threshold to determine the amount of predicted hot-water use that determines an actual heating phase. In this setting the eco-parameter is set to the respective cooling threshold $\beta = \Theta_c$. The output of the heating decision maker are tuples $(t, \Delta T)^n$ of the time as hour and the ΔT as positive temperature gradient over a period of n hours. The heating schedule is set up with a heating phase every time a $\Delta T > 0$ is present at the respective hour. The tuples are forwarded as input for the model 2 to predict the consumed energy. The output of model 2 are tuples $(t, m_{\text{hour}})^n$ with the time as hour and the summed modulation m_{hour} in that hour over the training or testing period of n hours. The energy consumption over n hours can be then calculated as

$$E = \sum_{i=1}^{n} m_{\text{hour}} \cdot c_{\text{heater}} \quad [\text{kWh}], \tag{3}$$

with m_{hour} as modulation per hour in percent and c_{heater} as the capacity of the domestic heater. In this work, we consider always $c_{\text{heater}} = 19$ kW.

B. Data Pre-processing and Feature Selection

In a first step, the state of the storage charge pump is used to distinguish the use of energy for space heating and for hot water heating. Every data example in which this pump is activated, the energy is counted to hot water preparation.
The data is present as data examples at a distance of one minute. It is almost complete in all data sets forming an suitable learning basis. Incomplete data tuples were deleted and not interpolated to prevent wrong decisions. It is not reasonable to attempt to make predictions of hot water consumption for each single minute. Instead, regular hourly patterns are searched. For this reason, the tuples that are used to compute the temperature gradients temperature were aggregated doing the average over 60 minute-intervals. The data was split up in one year training and a half year testing phase, resulting to 8760 training and 4368 testing examples.

After the data pre-processing, the following features were used to train both models:

- Unix timestamp in minutes used as index, as integer.
- outside temperature in °C as float.
- actual temperature in the water reservoir in °C.
- weekday derived from the timestamp as integer from 1 to 7.
- minutes of the actual day as integer from 0 to 1339.
- the actual season of the data example calculated from the timestamp and adapted to the periods of the northern hemisphere (EU) as integer from 0 to 3.
- modulation as power of the domestic heating system in percent as integer from 0 to 100.
- temperature gradient of two successive cylinder temperature data points from the first, second and third past entry in °C as float.
- modulation from the first, second and third past entry as integer from 0 to 100.

Recurrent features in this context include the values of the past n previous data entries (in this application up to three previous data entries) and creates so a connection between them. This is meaningful because a temperature trend is not only dependent on the actual status but often also from the previous states. Recurrent features are also a common approach for person re-identification with machine learning, where it brings more accuracy by training with previous frames of a person [13].

TABLE I
RESULTS OF NEURAL NETWORKS

RMSE	model 1		model 2	
	train	test	train	test
average	5.81 %	7.18%	3.23 %	7.28 %
best	1.89 %	3.21 %	2.10 %	3.04 %
worst	7.04 %	10.17 %	3.43 %	9.79 %

TABLE II
RESULTS OF GAUSSIAN PROCESSES

RMSE	model 1		model 2	
	train	test	train	test
average	4.20 %	7.22 %	1.97 %	7.71 %
best	2.84 %	4.85 %	3.03 %	4.50 %
worst	6.83 %	12.05 %	0.16 %	11.32 %

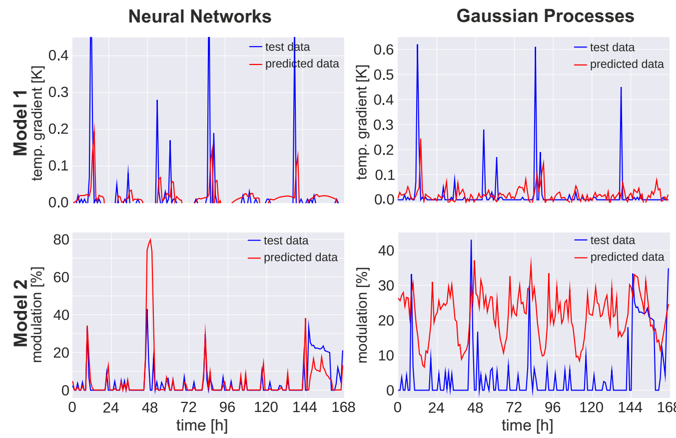

Fig. 3. Data vs. predictions of Neural Networks and Gaussian Process models.

C. Structure of the Models

We consider two different machine learning models using the same data and input features: Gausian process models and neural networks

1) Gaussian Process (GP) model: The GP-regression is performed using a standard radial basis function as kernel with an additional periodic (sinousoidal) term to better model the typical periodic patterns in water consumption. We set the mean function $m(x)$ to zero and the lengthscale γ to one because a priori not enough knowledge is available to define certain values.

$$m(x) = 0, \tag{4}$$

$$cov(x, \gamma) = \exp(\frac{-(x-x')^2}{2\gamma^2}) + \exp(-\frac{2\sin^2(\frac{x-x'}{2})}{\gamma^2}), \tag{5}$$

2) Neural Network model: We used a fully-connected multi-layer feed-forward network consisting of five hidden layers with different number of neurons with amounts of [500, 500, 200, 50, 5]. The design of the network was performed empirically; increasing the layers or neurons brought no significant improvement or tend to over-fitting. For the activation function we used the ReLU-function and the Adam optimizer with the Keras standard learning rate $\alpha = 0.001$ to train the network due to its good performance[14]. The batch size we used is 5 and the amount of epochs is 100. Regularization techniques are not used.

IV. RESULTS AND DISCUSSION

For both models and two algorithms the RMSE was calculated and reproduced in Table I and II.

The results for Neural Networks and Gaussian Processes are similar, the average test errors of the both models are all slightly smaller for neural networks. There is also a high variability in the data sets which is shown by the large distance between the best and worst error rates. Notable is also that best and worst error rates are calculated in different data sets, so there is not a user whose behavior could be learned at best by one of the two algorithms.

With a look at the predicted data in Fig. 3, we see that there are some large peaks in one week witch were not exactly matched in their amplitude by the algorithms. Notable is the prediction of Gaussian Process in model 2, here the values are wrongly predicted to high values for the whole period, because this week distinguishes from the other trained weeks too much and the periodic kernel component can not react on that, the neural networks act correctly on this change.

A. Energy Saving

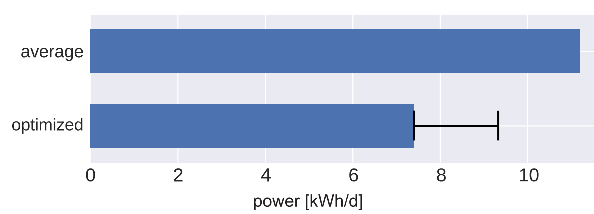

Fig. 4. Energy Savings with Error Interval.

Using the learned behavior with respect to the water extractions done by users, the heating times of the water heater can be adapted to a much smaller period compared to the default heating times so that water is only heated when it is needed. Fig. 4 shows in the upper bar the average daily consumption of the data sets within the testing phase during the winter months (from October till March) and the optimized average consumption in the lower bar during the same period with an optimized schedule computed based on the neural network predictions. The energy was calculated as described with Eq. 3 based on the standard heating schedule or the optimized. In both cases, the real energy consumption for the times when water was used outside of the computed schedule was added to the energy consumption computation. Whenever this happens, the user has to wait a few seconds, because the heater has to prepare hot water in this time, decreasing the achieved comfort.

The diagram states an average consumption without optimization of 11.20 kWh per day and a consumption with optimization of 7.41 kWh which corresponds to 66.15% of the original consumption. Taken the errors for the predictions in consideration, energy savings of 19.39% or 2.17 kWh per day can be expected in the worst case. The error interval (black-colored in Fig. 4) composes of the sum of the average RMSEs for model 1 and 2 as estimation for this scenario, because the data is processed through these two models.

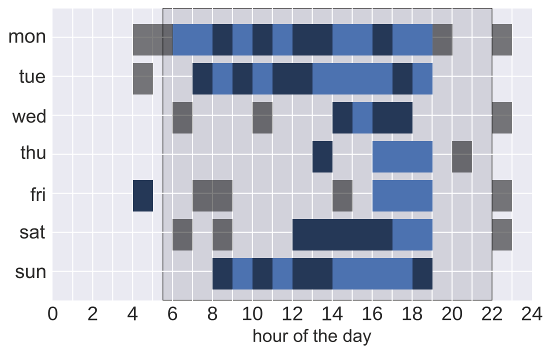

Fig. 5. Optimal heating schedule (blue bars) in comparison to the actual user behavior (grey bars) and to the standard heating time (grey shaded region).

The optimized heating schedule that leads to such energy savings is shown in Fig. 5 exemplary for the first week of the testing data and one data set / one user. The optimal heating is presented as blue bars. The grey bars show the actual heating schedule. Due to the testing prediction error of around 7 % not all times are matched with the predictions and so not all user behavior is included in the optimized heating schedule, but it shows a better adaptation than the standard heating times (the darker shaded region in the back). The optimized heating times (blue bars) are derived from the predictions with an eco-parameter $\beta = \Theta_c$.

B. Consumption of Resources

The training of the models differs in the consumption of different resources. In general, Gaussian Processes expend a huge amount of calculation time and memory more than neural networks. For the executions we used a Windows-PC with 2.4GHz QuadCore-CPU and 32 GB of RAM, furthermore we used a Python2.7 environment with the frameworks Keras and Tensorflow for neural networks and GPy for Gaussian Processes. The training phase of neural networks for both models takes time around 60 to 100 seconds and needs around 70 to 85 MB memory for loading the models. In contrast, Gaussian Processes need around 25 GB of RAM and took 20 to 40 minutes for the training of one model. The computation lasts significantly longer because of the matrix decomposition that needs to be performed for Gaussian Processes.

V. Conclusion

We have presented an approach to reduce the large amount of data obtained from IoT devices in a smart home setting. Real data from water heaters has been used to learn the human behavior related to domestic hot water usage. Models based on Gaussian processes and neural networks have been developed to compute optimized heating schedules that lead to significant energy savings in average for all considered users over a testing period of six months. Neural Networks deliver better results with smaller computational resources. Due to the considered prediction error of around 7% we were not able to predict all times a user requests hot water, but the optimized schedule shows a good approximation leading to significant energy savings.
In the future work this approach should be investigated with other machine learning algorithms like LSTMs for their suitability to reduce further the prediction error rate. The prepared optimization should be tested in an experimental plant to validate the results on a real device and replace model 2 with real energy measurements. In addition, we plan to integrate the presented method in smart energy management systems [15], [16] to further increase the energy savings in smart buildings.

References

[1] L. Perez-Lombard, J. Ortiz, and C. Pout, “A review on buildings energy consumption information,” *Energy and Buildings*, vol. 40, pp. 394–398, 01 2008.

[2] R. Saidur, H. Masjuki, and M. Jamaluddin, “An application of energy and exergy analysis in residential sector of malaysia,” *Energy Policy*, vol. 35, pp. 1050–1063, 02 2007.

[3] Eurostat (European Commission), *Energy balance sheets – 2015 data – 2017 edition*. EU publications, 7 2017.

[4] A. Delorme-Costil and J.-J. Bezian, “Forecasting domestic hot water demand in residential house using artificial neural networks,” *16th IEEE International Conference on Machine Learning and Applications*, p. 467472, 2017.

[5] M. C. Mozer, “The neural network house: An environment that adapts to its inhabitants,” in *AAAI Technical Report SS-98-02*. AAAI Technical Report SS-98-02, 1998, pp. 110–114.

[6] M. Aydinalp, V. I. Ugursal, and A. S. Fung, “Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks,” *Applied Energy*, vol. 79, no. 2, pp. 159 – 178, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261903002344

[7] V. Bakker, A. Molderink, J. L. Hurink, and G. J. M. Smit, “Domestic heat demand prediction using neural networks,” *2008 19th International Conference on Systems Engineering*, pp. 189–194, 2008.

[8] A. Rogers, S. Maleki, S. Ghosh, and J. Nicholas R, “Adaptive home heating control through gaussian process prediction and mathematical programming,” in *Second International Workshop on Agent Technology for Energy Systems (ATES 2011)*, May 2011, pp. 71–78, event Dates: May 2011. [Online]. Available: https://eprints.soton.ac.uk/272235/

[9] F. M. Gray and M. Schmidt, “Thermal building modelling using gaussian processes,” *Energy and Buildings*, vol. 119, pp. 119 – 128, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778816300494

[10] O. G. Santin, “Behavioural patterns and user profiles related to energy consumption for heating,” *Energy and Buildings*, vol. 43, no. 10, pp. 2662–2672, 2011.

[11] T. Hong, S. C. Taylor-Lange, S. DOca, D. Yan, and S. P. Corgnati, “Advances in research and applications of energy-related occupant behavior in buildings,” *Energy and Buildings*, vol. 116, pp. 694–702, 2016.

[12] S. Lucia, M. Kögel, P. Zometa, D. Quevedo, and R. Findeisen, “Predictive control, embedded cyberphysical systems and systems of systems–a perspective,” *Annual Reviews in Control*, vol. 41, pp. 193–207, 2016.

[13] Y. Yan, B. Ni, Z. Song, C. Ma, Y. Yan, and X. Yang, “Person re-identification via recurrent feature aggregation,” *CoRR*, vol. abs/1701.06351, 2017. [Online]. Available: http://arxiv.org/abs/1701.06351

[14] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *CoRR*, vol. abs/1412.6980, 2014. [Online]. Available: http://arxiv.org/abs/1412.6980
[15] B. Karg and S. Lucia, “Towards low-energy, low-cost and high-performance iot-based operation of interconnected systems,” in *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*. IEEE, 2018, pp. 706–711.
[16] ——, “Deep learning-based embedded mixed-integer model predictive control,” in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 2075–2080.