ACTIVITY-BASED RECOMMENDATIONS FOR THE REDUCTION OF CO2 EMISSIONS IN PRIVATE HOUSEHOLDS

Alona Zharova*
Humboldt-Universität zu Berlin
Berlin, Germany
alona.zharova@hu-berlin.de

Laura Löschmann
Humboldt-Universität zu Berlin
Berlin, Germany
lauraloeschmann@gmx.net

ABSTRACT

This paper proposes an activity prediction framework for a multi-agent recommendation system to tackle the energy-efficiency problem in residential buildings. Our system generates an activity-shifting schedule based on the social practices from the users' domestic life. We further provide a utility option for the recommender system to focus on saving CO2 emissions or energy costs, or both. The empirical results show that while focusing on the reduction of CO2 emissions, the system provides an average of 12% of emission savings and 7% of electricity cost savings. When concentrating on energy costs, 6% of emission savings and 20% of electricity cost savings are possible for the studied households.

1 INTRODUCTION

The energy consumption of private households amounts to approximately 30% of the total global energy consumption (Allouhi et al., 2015), causing a large share of the CO2 emissions through energy production. Increasing the efficiency of energy consumption through managing the demand, for instance through load shifting, is a feasible way to reduce CO2 emissions. For the practical implementation of load shifting, private households require data-driven decision support. Recommendation systems provide a framework by suggesting energy-efficient actions.

Existing literature on recommender systems for load shifting in residential houses mainly focuses on rescheduling the device usage (Marinakis & Doukas, 2018; Khalid et al., 2019; Fioretto et al., 2017). However, recommending residents when to use appliances does not resonate with the way people spend their time at home. The domestic use of energy is a result of deeply embedded social practices (Katzeff & Wangel, 2015), which can be depicted as activities (i.e., cooking, laundering). Dwellers can more easily understand and follow the recommendations on shifting their activities instead of devices (Stankovic et al., 2016). Only a few works explore activity-aware systems and recommend the shifting of domestic activities corresponding to energy-consuming devices (Marcello & Pilloni, 2020; Thomas & Cook, 2016). However, these works rely on various sensor data with activity labels. The high amount of sensors increases the implementation effort and costs in real-life applications and burdens users with the requirement of tracking their activities for a sufficient amount of time.

This research focuses on human-centered energy-efficiency improvements in residential households to tackle climate change. In particular, we propose an activity prediction framework for a multi-agent recommendation system to nudge behavioral changes. We summarize our key contributions as follows. First, we design and implement an Activity Agent that calculates the activity probabilities for every hour in contrast to other works that utilize daily device-usage probabilities (Jiménez-Bravo et al., 2019; Riabchuk et al., 2022; Zharova et al., 2022). The resulting activity-shifting schedule is based on social practices from the users' domestic life. Therefore, recommendations can be easier integrated into daily life, fostering the acceptance and usage of the system over a longer period. Second, aiming at practical deployment, we suggest a measure to evaluate the Activity Agent's performance without ground truth data. Third, we enhance the prediction power of the system by utilizing Random Forest (RF) and feed-forward Neural Network (NN) algorithms in comparison to

*Corresponding author.

the benchmark model in Riabchuk et al. (2022). Fourth, to produce recommendations, our system needs minimal user input, and appliance-level energy consumption data (i.e., from low-cost smart plugs) and does not require activity labels or other sensors, thus, reducing the implementation costs. With this, the system solves the cold start problem in two to six months for most households. Finally, we provide a utility option to focus on saving CO2 emissions or energy costs, or both.

2 Methods

Our activity-based recommendation system builds upon the architecture of [9] that includes six agents: Price Agent (collects external electricity price data), Preprocessing Agent (prepares the appliance energy consumption data), Load Agent (generates device load profiles), Availability Agent (produces hourly probabilities of user availability), Usage Agent (calculates daily device usage probabilities), and Recommendation Agent (generates a device-shifting schedule for the next 24 hours). However, the recommendation system by Riabchuk et al. (2022) generates device-usage recommendations without considering activities and, thus, does not resonate with the social practices of domestic life.

In this work, we propose an activity prediction framework for the multi-agent recommendation system (see Figure 1). First, we divide the activities into three groups referring to their flexibility to shift starting hours: flexible, slightly flexible, and inflexible. Flexible activities can be easily brought forward or postponed to later (i.e., cleaning), slightly flexible have a certain potential to be shifted but not throughout the whole recommendation horizon (i.e., entertaining), and inflexible are bound to certain starting hours (i.e., working). During the installation of the system, the user has to specify which household devices correspond to which activities. Using the user input, the system creates an activity-device mapping in a form of a vector representation. Second, we repurpose the Usage Agent to calculate the hourly usage probabilities for all devices related to flexible activities. Third, we create an Activity Agent that receives the output of the Usage Agent and the activity-device mapping. It applies a vector space model to calculate the cosine similarity for each activity vector and outputs the hourly probabilities for every flexible activity for the 24-hour recommendation horizon. Next, to address the CO2 reduction within the recommender system, we create a CO2 Emissions Agent that predicts the hourly amount of CO2 emissions to be generated during the power production in the recommendation horizon. The Recommendation Agent collects outputs from other agents and generates activity-shifting recommendations.

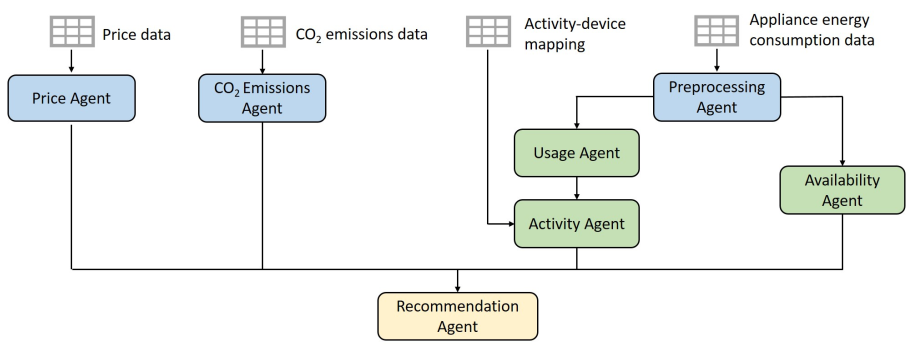

Figure 1: Architecture of the activity-based multi-agent recommendation system

The Availability Agent takes five features as input: month, day and hour of appliance use, the resident's availability one hour and one week ago based on the use of appliances that ensure the resident is at home. The Usage Agent takes the time features (month, day, hour) and time lag one-hot-encoded features (one hour and one week ago) indicating the usage of each household appliance. Since the Availability and the Usage Agents perform a classification task with supervised learning, we measure their performance using an Area Under the ROC-Curve (AUC). Initially, the models within the agents are trained on the four weeks of data to address the cold start problem. We calculate the AUC using the fifth week as a test set. With each passing day, the dataset gets larger, and so does the training and test set, maintaining an 80:20 ratio. After every four months, the hyperparameters are tuned cutting of the oldest two months to ensure the model accounts for the household's current consumption behavior. In this way, the models are trained using a maximum of six months of energy consumption data. The final AUC is the average AUC value of all evaluations within the one year. To prevent overfitting due to the small dataset, we apply simple model and early stopping.

The Activity Agent utilizes unsupervised learning. Following our middle-term aim of practical deployment of the system with minimal user input, we develop an EQUAL score to evaluate the Activity Agent's performance without ground truth data (see Appendix for derivation details). The EQUAL score indicates how many hourly sets of activities out of 24 hourly sets the Activity Agent can correctly predict on average. To evaluate the overall performance of the recommendation system, we extend the evaluation framework of Riabchuk et al. (2022) to account for CO2 emission savings. In particular, we quantify how much CO2 emissions and costs a user can save by employing the system for one year. For this purpose, we calculate the CO2 emissions and energy costs for executing the activities with and without following the activity-shifting recommendations generated by the system.

3 EMPIRICAL RESULTS

To analyze the proposed recommendation system, we use the REFIT dataset (Murray et al., 2017) that contains appliance-level electrical consumption data of 20 households in the UK from 2013 to 2015. Additionally, we use the carbon intensity data (ESO) and the day-ahead electricity prices for the UK (entsoe). For our empirical study, we manually associate the devices with the corresponding activities for every household to prepare the activity-device mappings. The recommendation system further requires the value of the emissions ratio from the user as an input (i.e., 1.0 in case of focusing on emission savings, 0.0 for cost saving, or 0.5 for both equally).

The recommendation system has three more hyperparameters: an availability threshold, an activity threshold, and an *aval-off* value. The availability threshold indicates the minimum predicted probability that the user will be available, the activity threshold means the minimum predicted probability that an activity will be carried out, and the Boolean variable *aval-off* means turning off the dependence on the predicted availability hours for recommendations. These three hyperparameters are initialized by the system, however, can be adjusted by the user later. Only hours with the probability of user availability and activity higher than a threshold are considered for a recommendation.

Sintov & Schultz (2017) show that adjustable green defaults maximize energy savings. Therefore, we initialize the system with hyperparameter values that lead to the highest possible savings. For optimal initialization, we analyze changes in the distribution of the recommended activity launching hours as well as in the performance measures due to changes in hyperparameters. We also perform a grid search for the highest total savings on the household level to confirm the best parameters were found. As a result, the availability and activity thresholds are set to 0.15 and the *aval-off* value to true.

Riabchuk et al. (2022) use Logistic Regression for the prediction tasks of the Availability and the Usage Agents. To improve the prediction performance in comparison to other studies (Riabchuk et al., 2022; Zharova et al., 2022), we train the RF and NN models on the first four weeks of data for each household before they make their first prediction. The models are tuned every four months using the new energy consumption data and cutting off the oldest two months of data to focus on the current energy consumption behavior. The NN model outperforms and, therefore, is used further. We calculate the EQUAL score for the Activity Agent for various households (see Table A4 in Appendix for more details). For instance, the EQUAL score of 0.79 for household 5 indicates that the Activity Agent predicts 19 out of 24 hourly sets of activities on average correctly.

To measure the impact of the cold start problem on the recommender system, we evaluate the Availability, Usage, and Activity Agents for varying lengths of data. The Availability Agent reaches rapidly acceptable performance solving the cold start problem after about 2.5 months, which is relatively constant across households. On the contrary, the cold start scores of the Usage Agent vary across the different devices and households. The cold start problem of the Activity Agent is solved after around one month which is also quite constant across the households. The recommendation system solves the cold start problem when all agents solve it. For instance, for household 5 all three agents solve the cold start problem in 98 days at the earliest. Thus, the complete framework is solved in 98 days (see Table A5 in Appendix).

Table 2: Performance of the recommendation system for households 1 to 5 with an emissions ratio of 0.0 compared to an emissions ratio of 1.0 and constant availability and activity thresholds of 0.15

			emissions ratio = 0.0				emissions ratio = 1.0			
			emissions sav.		price sav.		emissions sav.		price sav.	
HH	#recom.	#recom./day	total	relative	total	relative	total	relative	total	relative
1	1,837	5	123.37	0.07	79.84	0.22	261.32	0.14	21.65	0.06
2	3,231	9	188.35	0.07	109.54	0.20	345.72	0.13	40.45	0.07
3	3,084	8	184.85	0.06	119.63	0.20	355.63	0.12	32.66	0.06
4	2,216	6	30.43	0.03	27.76	0.13	80.15	0.07	-0.38	0.00
5	3,893	11	219.30	0.07	136.46	0.25	350.97	0.13	74.77	0.14

The empirical results show that the average day-ahead CO2 emissions per hour and the average day-ahead prices per hour are positively correlated (see Figure A2). Therefore, focusing on emission savings only can also lead to cost savings and vice versa. Further analysis of different values for the activity and the availability thresholds (see Tables A6 - A9) shows that the highest possible savings are achieved with the low availability threshold combined with the low activity threshold. Higher values of the activity and the availability thresholds lead to a reduction of the number of recommendations and, in turn, a reduction of possible savings, but a similar distribution of recommendation timings (see Figures A3 - A6). The empirical results in Table 2 indicate that while focusing on CO2 emissions (emissions ratio = 1.0), the system provides an average of 12% of CO2 emission savings and 7% of electricity cost savings. When setting the focus on energy cost savings (emissions ratio = 0.0), 6% of emission savings and 20% of electricity cost savings on average are possible for the studied households in case of acceptance of the recommendations.

4 Discussion

The proposed recommendation system draws on the previous works by using simple hardware, the flexible architecture of a multi-agent system as well as a similar logic to generate recommendations. However, our system improves and enhances the previous findings in several aspects.

(i) The recommendation systems by Riabchuk et al. (2022) and Jiménez-Bravo et al. (2019) generate a maximum of one recommendation per device per day. Our system can suggest one recommendation per activity per hour. In other words, our recommendation system provides one recommendation for every time the probability for an activity to happen exceeds the activity threshold in the recommended time horizon. This reflects the reality since domestic activities such as cooking or laundering can happen several times a day.

(ii) Riabchuk et al. (2022) focus on shifting the usage of flexible devices (i.e., dishwasher), Jiménez-Bravo et al. (2019) cover all shiftable devices but limit the number of recommendations per device to not overload the user with too many recommendations. Recommending the shifting of activities reduces the number of recommendations per se, since activities represent a specific number of devices that are used to carry out the activity.

(iii) The recommendation systems of Riabchuk et al. (2022) and Jiménez-Bravo et al. (2019) aim at saving energy costs with load shifting. The proposed system enhances the utility dimension by the possibility of setting the focus on saving CO2 emissions or electricity costs, or both equally. Thereby, the system reaches a wider target group. In addition, the user can switch the focus of the system while using it.

5 Impact and scaling potential

A large fraction of CO2 emissions in high-income countries is due to energy consumption in residential buildings. In low- and middle-income countries, this share is even higher. Suggesting activity-shifting recommendations to private households every day provides a guide to control their own ecological footprint and energy costs and furthermore supports SGD 7, 9, 11, and 13. Our system can be used within a smartphone application or existing smart home system, thus, being a viable tool for climate change mitigation in low-, middle-, and high-income countries alike.

We see the following scaling potential for our approach. In 2023 20% of private households in Germany will have a smart home system aiming at energy management (Statista, 2021). If 5% of these households utilize the proposed activity-based multi-agent recommendation system, then 32 kilotons of CO_2 emissions could be saved in 2023 in Germany. In 2026 the saving potential is even higher since almost 48% of German households would be using smart home systems for energy management (see Appendix for calculation details).

6 Conclusions

The empirical results show that our solution fosters saving CO_2 emissions and energy costs by shifting domestic activities. Activity-based recommendations can be more easily integrated into daily life. This facilitates the acceptance and the usage of the system in a long-term perspective and, thus, tackling the energy-efficiency problem in residential households.

Acknowledgments

We would like to thank Elie Alhajjar for his supportive mentorship within the CCAI program while preparing this paper for submission. We wish to express our sincere gratitude and warm appreciation to Stefan Lessmann for the insightful feedback that has greatly improved this paper.

REFERENCES

Amine Allouhi, Youness El Fouih, Tarik Kousksou, Abdelmajid Jamil, Youssef Zeraouli, and Youssef Mourad. Energy consumption and efficiency in buildings: current status and future trends. *Journal of Cleaner Production*, 109:118–130, 2015.

Destatis. Statistisches Bundesamt. Entwicklung der Privathaushalte bis 2040, 2020. URL `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Publikationen/Downloads-Haushalte/entwicklung-privathaushalte-5124001209004.pdf?__blob=publicationFile`.

entsoe. Transparency platform. day-ahead prices [online database]. URL `https://transparency.entsoe.eu/transmission-domain/r2/dayAheadPrices/show`.

National Grid ESO. Carbon intensity api [online database]. URL `https://www.carbonintensity.org.uk/`.

Ferdinando Fioretto, William Yeoh, and Enrico Pontelli. A multiagent system approach to scheduling devices in smart homes. AAMAS '17, pp. 981–989, Richland, SC, 2017.

Petra Icha, Thomas Lauf, and Gunter Kuhs. Entwicklung der spezifischen Kohlendioxid - Emissionen des deutschen Strommix in den Jahren 1990 - 2020. Umweltbundesamt. *Climate Change*, 15 (2022), 2022. URL `https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2022-04-13_cc_15-2022_strommix_2022_fin_bf.pdf`.

Diego M. Jiménez-Bravo, Javier Pérez-Marcos, Daniel H. De la Iglesia, Gabriel Villarrubia González, and Juan F. De Paz. Multi-agent recommendation system for electrical energy optimization and cost saving in smart homes. *Energies*, 12(7), 2019.

Cecilia Katzeff and Josefin Wangel. Social practices, households, and design in the smart grid. In *ICT Innovations for Sustainability*, Advances in Intelligent Systems and Computing, pp. 351–365. Springer, 2015.

Rabiya Khalid, Nadeem Javaid, Muhammad Hassan Rahim, Sheraz Aslam, and Arshad Sher. Fuzzy energy management controller and scheduler for smart homes. *Sustainable Computing: Informatics and Systems*, 21:103–118, 2019.

Francesca Marcello and Virginia Pilloni. Smart building energy and comfort management based on sensor activity recognition and prediction. Eclipse Foundation, 2020.

Vangelis Marinakis and Haris Doukas. An advanced IoT-based system for intelligent energy management in buildings. *Sensors*, 18(2), 2018.

David Murray, Lina Stankovic, and Vladimir Stankovic. An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study. *Scientific Data*, 4(1):160122, 2017.

Valentyna Riabchuk, Leon Hagel, Felix Germaine, and Alona Zharova. Utility-based context-aware multi-agent recommendation system for energy efficiency in residential buildings. *arXiv:2205.02704*, 2022.

Nicole D. Sintov and Wesley P. Schultz. Adjustable green defaults can help make smart homes more sustainable. *Sustainability*, 9(4), 2017.

Lina Stankovic, Vladimir Stankovic, Jing Liao, and Clevo Wilson. Measuring the energy intensity of domestic activities from smart meter data. *Applied Energy*, 183:1565–1580, 2016.

Statista. Research department. Prognose zur Anzahl der Smart Home Haushalte nach Segmenten in Deutschland für die Jahre 2017 bis 2025 (in Millionen). [Infographic], 2021. URL `https://de.statista.com/prognosen/801573/anzahl-der-smart-home-haushalte-nach-segmenten-in-deutschland`.

Statista. Research department. Jährlicher Stromverbrauch eines privaten Haushaltes in Deutschland in den Jahren 1991 bis 2021 (in Kilowattstunden). [Infographic], 2022. URL https://de.statista.com/statistik/daten/studie/245790/umfrage/stromverbrauch-eines-privathaushalts-in-deutschland/.

Brian L. Thomas and Diane J. Cook. Activity-aware energy-efficient automation of smart buildings. *Energies*, 9(8), 2016.

Alona Zharova, Annika Boer, Julia Knoblauch, Kai Ingo Schewina, and Jana Vihs. Explainable multi-agent recommendation system for energy-efficient decision support in smart homes. *arXiv:2210.11218*, 2022.

A Appendix

Table A1: An example of an activity-device mapping in form of a vector representation

Device / Activity	Washing machine	Stove	Television	Kettle	Dishwasher
Laundering	1	0	0	0	0
Cooking	0	1	0	1	0
Entertaining	0	0	1	0	0
Cleaning	0	0	0	0	1

Table A2: An exemplary output of the recommendation system

Recommendation Date:	29/08/2021			
Greenest Hour of the Day:	8			
Cheapest Hour of the Day:	4			
Activity	Availability Flag	Activity Flag	Best Hour	Duration
Cooking	0	0	11	2
	0	0	18	1
Laundering	1	0	4	4
Entertainment	1	0	-	-
Working	0	0	10	4
Total emission savings:	in kgCO_2/kWh			
Total price savings:	in €			

The output of the system starts with an overview of the greenest hour, meaning the hour with the fewest carbon intensity forecast of the day, and the cheapest hour of the day not considering predicted availability hours. The system then generates the best beginning hours combined with the predicted activities' duration. The system concludes its output by providing the possible emissions and price savings for the recommended date in reference to the emissions ratio specified by the user that are achievable by implementing all provided recommendations. Based on this data the recommendation system provides activities' beginning hours for the next 24 hours starting from the point in time the recommendations are provided by the system. The recommendations are made if the user's availability is predicted by the system and the suggested activity schedule would reduce energy costs or CO2 emissions.

Table A3: Performance evaluation of the tested models Logistic Regression (LR), Random Forest (RF) and Neural Network (NN) for the individual agents using the data of household 5

Model	AUC Availability	AUC Usage								EQUAL Activity
		Dev1	Dev2	Dev3	Dev4	Dev5	Dev6	Dev7	Dev8	
LR	0.75	0.90	0.88	0.84	0.89	0.94	0.96	0.98	0.91	0.76
RF	0.83	0.84	0.82	0.67	0.89	0.94	0.91	0.95	0.90	0.75
NN	0.87	0.95	0.91	0.88	0.90	0.95	0.96	0.99	0.89	0.79

Table A4: Performance evaluation of the individual agents for the households (HH) 1 to 5. The Neural Network model used for the predictions of the Availability and the Usage Agents

HH	AUC Availability	AUC Usage								EQUAL Activity
		Dev1	Dev2	Dev3	Dev4	Dev5	Dev6	Dev7	Dev8	
1	0.87	0.74	0.82	0.78	0.86	1.00	-	-	-	0.46
2	0.84	0.92	0.94	0.98	0.93	0.89	0.99	0.98	-	0.57
3	0.78	0.96	0.95	0.89	0.95	0.80	0.88	1.00	-	0.69
4	0.84	0.82	0.98	0.95	0.95	0.99	1.00	-	-	0.64
5	0.87	0.95	0.91	0.88	0.90	0.95	0.96	0.99	0.89	0.79

Table A5: Cold start scores in days for the households 1 to 5

HH	Availability	Usage								Activity	System
		Dev1	Dev2	Dev3	Dev4	Dev5	Dev6	Dev7	Dev8		
1	81	365	74	80	103	28	-	-	-	51	365
2	72	28	41	30	28	67	28	28	-	35	71
3	73	99	33	29	28	28	120	29	-	37	120
4	72	186	75	84	78	32	29	-	-	32	186
5	78	28	28	98	28	28	28	28	28	30	98

Table A6: Changes in the performance measures for household 5 regarding different activity thresholds, an emissions ratio of 1.0 and a constant availability threshold of 0.15

			emissions savings		price savings	
activity th.	#recom.	#recom./day	total	relative	total	relative
0.15	3,893	11	350.97	0.13	74.77	0.14
0.40	3,088	8	139.14	0.12	23.92	0.10
0.65	1,655	5	53.08	0.13	7.09	0.10

Table A7: Changes in the performance measures for household 5 regarding different activity thresholds, an emissions ratio of 0.0 and a constant availability threshold of 0.15

			emissions savings		price savings	
activity th.	#recom.	#recom./day	total	relative	total	relative
0.15	3,893	11	219.30	0.08	136.46	0.25
0.40	3,088	8	86.38	0.07	50.10	0.22
0.65	1,655	5	29.68	0.07	17.77	0.25

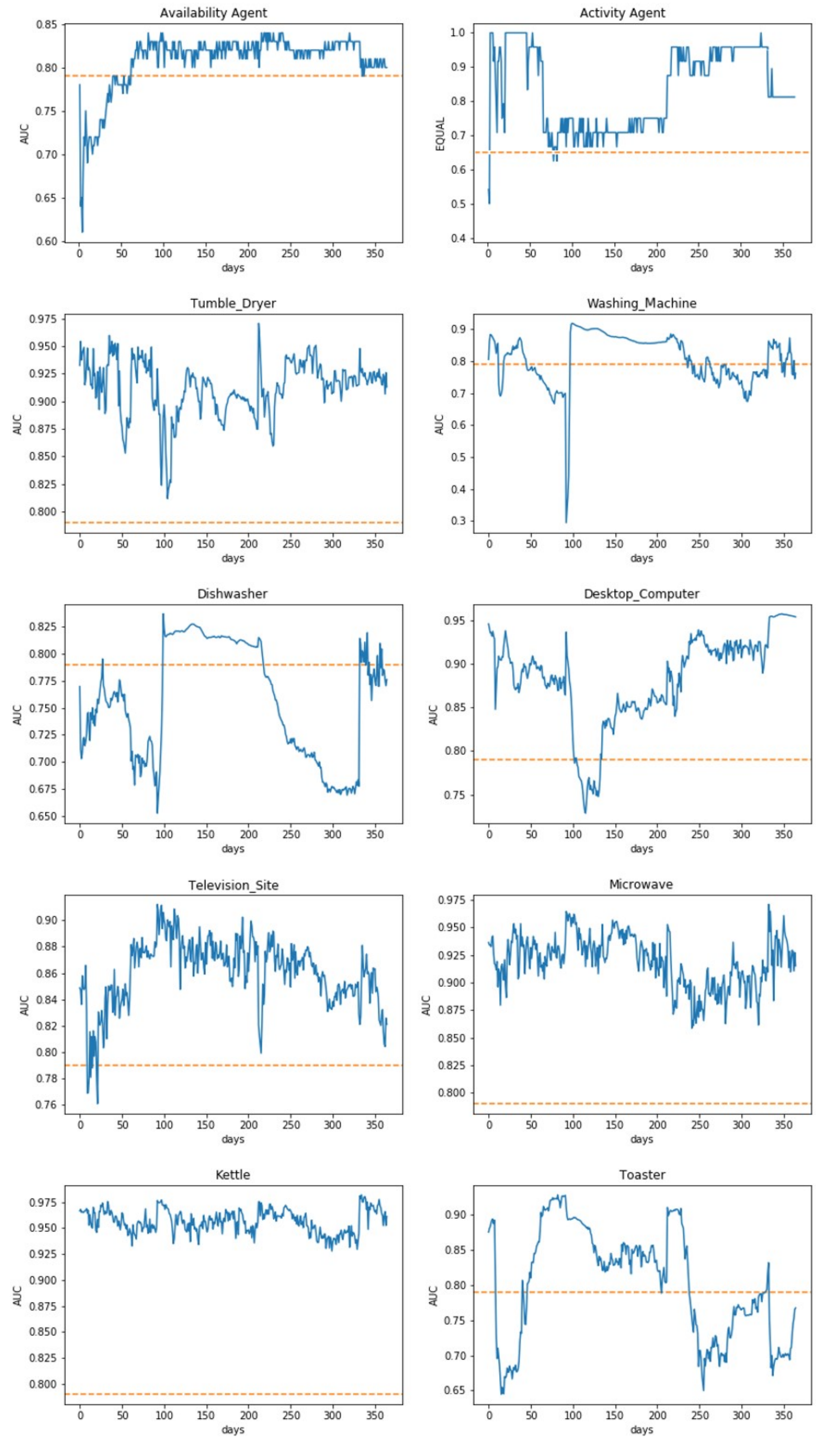

Figure A1: Performance of the Availability Agent, the Activity Agent and the Usage Agent over a year for household 5 for the cold start problem evaluation

Table A8: Changes in the performance measures for household 5 regarding different availability and activity thresholds, and an emissions ratio of 0.0

				emissions savings		price savings	
availability th	activity th	#recom.	#recom. /day	total	relative	total	relative
0.40	0.15	2,614	7	195.66	0.10	117.23	0.31
0.65	0.15	1,501	4	134.84	0.11	76.23	0.33
0.25	0.25	3,051	8	154.94	0.08	91.87	0.26

Table A9: Changes in the performance measures for household 5 regarding different availability and activity thresholds, and an emissions ratio of 1.0

				emissions savings		price savings	
availability th	activity th	#recom.	#recom. /day	total	relative	total	relative
0.40	0.15	2,614	7	310.19	0.16	66.04	0.17
0.65	0.15	1,501	4	204.99	0.17	44.98	0.19
0.25	0.25	3,051	8	249.16	0.13	45.88	0.13

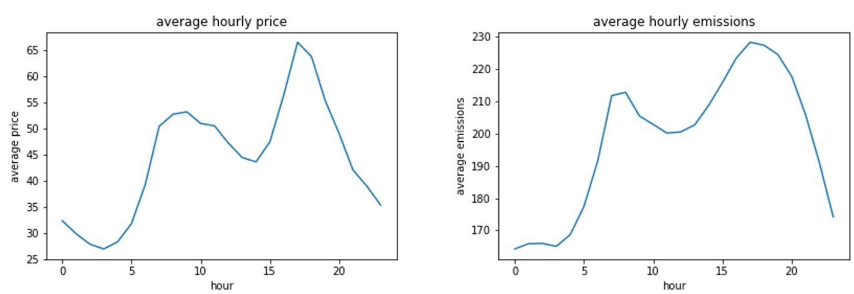

Figure A2: Average day-ahead hourly prices (in €/MWh) and average day-ahead hourly emissions (in gCO2/kWh)

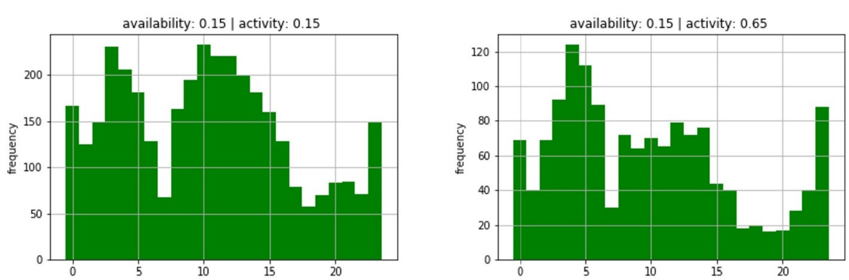

Figure A3: Changes in the distribution of starting hours for household 5 regarding different activity thresholds, an emissions ratio of 1.0 and a constant availability threshold of 0.15

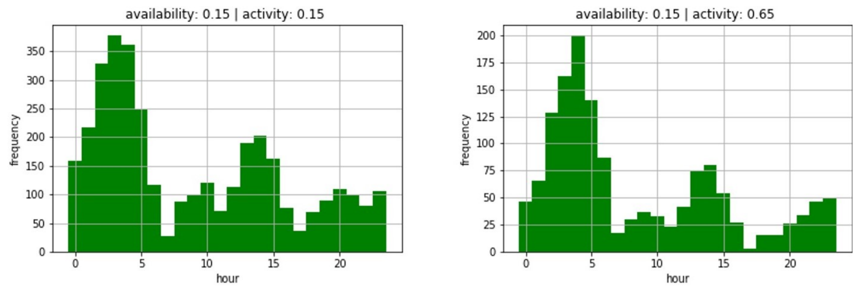

Figure A4: Changes in the distribution of starting hours for household 5 regarding different activity thresholds, an emissions ratio of 0.0 and a constant availability threshold of 0.15

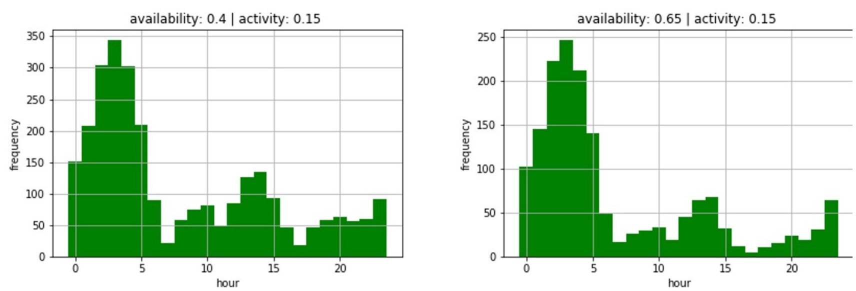

Figure A5: Changes in the distribution of starting hours for household 5 regarding different availability and activity thresholds, and an emissions ratio of 0.0

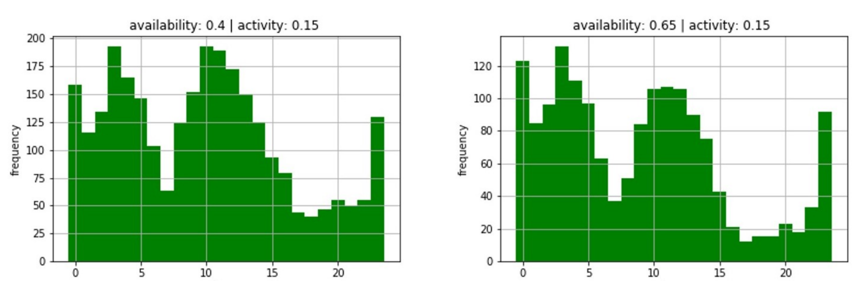

Figure A6: Changes in the distribution of starting hours for household 5 regarding different availability and activity thresholds, and an emissions ratio of 1.0

Table A10: Data for the calculation of the scaling potential

Variables		Values			Source
Short name	Full name	Fixed	2023	2026	
H	Number of German households		41738000	41791000	Destatis (2020)
H_{SH}	Number of German smart home households aimed at Energy Management		8500000	20000000	Statista (2021)
E_H	Electricity consumption per average household per year (kWh)	3050[1]			Statista (2022)
$CO2_E$	CO2 emissions per kWh of electricity (kg/kWh)	0.42[1]			Icha (2022)
	Possible CO2 emission savings (emissions ratio = 1.0)	11.8%[2]			
	Possible CO2 emission savings (emissions ratio = 0.0)	6%[2]			

1 - data for 2021
2 - 11.8% and 6% are the calculated averages of the relative CO2 emission savings based on our empirical results

Table A11: Calculation of the scaling potential

Variable	Description	Measure	Results 2023	Results 2026
$Ratio_{H_{SH}}$	Proportion of German smart home households aimed at Energy Management $(H_{SH} \times 100)/H$	%	**20.37**	**47.86**
$H_{SH_{5\%}}$	5% of German SH households aimed at Energy Management $(H_{SH} \times 5)/100$		425000	1000000
$E_{H_{SH_{5\%}}}$	Electricity consumption of 5% of smart home households aimed at Energy Management $H_{SH_{5\%}} \times E_H$	kWh	1296250000	3050000000
$CO2_{H_{SH_{5\%}}}$	CO2 emissions from 5% of SH multi-agent recommendation system $E_{H_{SH_{5\%}}} \times CO2_E$	kg	544425000	1281000000
$S_{CO2_{11.8\%}}$	CO2 savings for 11.8% of possible savings from multi-agent recommendation system $CO2_{H_{SH_{5\%}}} \times 11.8\%$	kg	64242150	151158000
$S_{CO2_{6\%}}$	CO2 savings for 6% of possible savings from multi-agent recommendation system $CO2_{H_{SH_{5\%}}} \times 6\%$	kg	**32665500**	**76860000**

The calculation results used in the main text are in bold.

B Appendix

Derivation of the EQUAL score

The predictions of the Activity Agent are based on the device usage probabilities. If a device's usage is likely in a certain hour, the activity that has this device as its identifying device is also likely to be carried out. Taking this principle every hour a set of devices S^i_{dev} that are predicted with a probability higher than a certain usage threshold can be translated in a set of activities S^i_{act} containing all activities, the devices in S^i_{dev} are identifying devices for. As a result, S^i_{act} can be used as a target variable to compare the predictions of the Activity Agent with.

The output of the Activity Agent are activity probabilities for every possible activity per hour. An hourly set of predicted activities $S^i_{\widehat{act}}$ can be compiled by taking all activities higher than a certain activity threshold. To measure the performance of the Activity Agent, both sets needs to be compared and checked for equality, meaning both containing the same activities. This procedure can be denoted in the following:

$$ID_{act} = \{dev \mid P(dev) = 1\}, \tag{1}$$

where ID_{act} is the set of identifying devices for each possible activity of the private household. The values for $P(dev)$ are provided by the activity-device mapping vector where 1 represents an identifying relation between the activity and the device.

$$S^i_{dev} = \{dev \mid \pi^{dev} > use_{th}\}, \tag{2}$$

where S^i_{dev} describes the set of devices whose device usage probability of hour i is greater than the usage threshold use_{th}.

$$S^i_{act} = \{S_{act} \mid dev \in S^i_{dev} \wedge \widehat{dev} \in ID_{act}\}, \tag{3}$$

where S^i_{act} is the set of activities that have an identifying activity-device relationship with the devices of S^i_{dev}.

$$S^i_{\widehat{act}} = \{\widehat{act} \mid \pi^{\widehat{act}} > act_{th}\}, \tag{4}$$

where $S^i_{\hat{act}}$ describes the set of activities that are predicted by the Activity Agent with a probability greater than the activity threshold act_{th}.

$$EQUAL_{act} = \sum_{i=1}^{n} (S^i_{act} = S^i_{\widehat{act}})/n, \tag{5}$$

where $EQUAL_{act}$ is the ratio of the sum of equal activity sets per hour i over n, with n being the total number of hourly activity sets in the recommendation horizon, in this case 24. Since the Usage Agent performs a preprocessing step for the Activity Agent also their performance evaluation is obviously linked. Therefore, for each prediction from the different models trained to evaluate the Usage Agent, the performance of the Activity Agent is measured using the proposed $EQUAL_{act}$ measurement. To receive a single evaluation value for the Activity Agent, the different measurements are combined by calculating their mean.