Proactive Smart Home Assistants - User's Characteristic-Based Preference Prediction with Machine Learning Techniques

Tianzhi He,[1] Farrokh Jazizadeh, Ph.D.[2]

[1]Graduate Research Assistant, Department of Civil and Environmental Engineering, Virginia Tech, 315 Patton Hall, 750 Drillfield, Blacksburg, VA, 24061; email: tianzhi@vt.edu
[2]Assistant Professor, Department of Civil and Environmental Engineering, Virginia Tech, 200 Patton Hall, 750 Drillfield, Blacksburg, VA, 24061; email: jazizade@vt.edu

ABSTRACT

AI-powered smart homes bring high-quality intelligent services to occupants with digital virtual assistants. Through interactions with occupants, the Smart Home Assistants (SHAs) can develop occupants' profiles using a number of personal characteristic features for tailored and smart interactions. Based on these profiles, smart home systems can *proactively* offer automation services while conserving occupants' comfort and convenience. In this study, we have sought to investigate characteristic features that affect occupants' perception of the proactive concept, as well as their preferences for modes of interactions through an application of automation for energy efficiency management. Upon a data collection through an online experiment on campus, 58 valid responses with personal characteristic features were utilized to develop predictive machine learning models. These models can predict participants' general attitude towards proactive SHAs, as well as their preferences for interaction modes with high performance (accuracy between 0.67 to 0.82 and F-score between 0.66 to 0.74). Various features were identified to have considerable significance, including personal beliefs about taking actions and energy expenses, as well as environmental protection values. The findings of this study provide an insight into the design of learning processes for virtual assistants in smart home ecosystems and the effect of the individual characteristics on the users' preferences for interactions with SHAs.

Key Words: Smart Homes, Home Automation, Machine Learning, Proactive Virtual Assistant, Energy Management

INTRODUCTION

In the past decade, the rapid development of Internet of Things (IoT) technologies enabled smart homes to improve occupants living quality. Built with a network that connects various smart home devices to exchange information and provide services, smart home systems with high level of automation can now change many aspects of the occupants' daily lives, such as control, convenience, comfort, and energy-saving (Alaa et al. 2017). Occupants usually interact with the smart home environment through an integrated interface that have the control of various systems and appliances (Mennicken et al. 2016). Commercial artificial intelligence (AI) virtual assistant products like Amazon Alexa, Google Assistant, and Apple Siri have been implemented in smart home ecosystems to provide occupants with convenient conversational agents. For example, through the Amazon Echo device, occupants can easily control broad compatible smart devices (e.g., smart lights and smart thermostats) with voice commands to the intelligent virtual assistant "Alexa".

Currently, most of the interactions between the occupants and smart home virtual assistants are initiated by users, i.e., most of these agents are reactive and can only respond to commands, which significantly limits the capabilities of smart homes (Miksik et al. 2020). However, the adoption of machine-learning functionalities and the applications of human-robot interactions bring new potentials in the form of proactive smart home virtual assistants for different applications including energy efficiency. The next generation of smart home assistants can not only enable occupants to have convenient control over the household appliances, but also learn the occupants' preferences based on their personal information and behave proactively for energy-saving and occupant's comfort (Jivani et al. 2018).

Previous studies have found that occupant's demographic backgrounds and personal characteristics such as values and beliefs can significantly affect their perception of smart home technologies and engagement with the smart home virtual assistants (Georgiev and Schlögl 2018; Tabassum et al. 2019). Nevertheless, limited studies have investigated occupants' perception of smart homes with high level of autonomy for energy efficiency, and the impact of occupants' characteristic features on being receptive to smart homes with proactive SHAs. To this end, a number of questions could be explored. How do occupants perceive the smart homes with automatic control and intelligent virtual assistants for energy efficiency? Can we move towards establishing occupants' profiles for effective communication between smart home assistants and occupants? What are the influential features in occupants' profiles that affect their interaction with the smart home assistants?

In this study, we conducted an online experiment using questionnaires to collect data from students and employees on campus and developed machine learning models (SVM, Random Forest, Logistic Regression) to move towards addressing the above-mentioned questions. This study provides an analysis of the effects of the individual characteristics on users' perception of proactive smart home ecosystems. The research methods, results and findings of this study will be further introduced in the following sections.

METHODOLOGY

Interactive Questionnaire Design for Online Experiment. To answer the research questions, a questionnaire was designed and distributed among the students and employees on campus. In the questionnaire, we created an experimental setting in which the smart home assistants have the automatic control over the thermostat, with the goal of helping occupants save energy. Therefore, participants were introduced to SHAs increased autonomy in the context of automation for energy efficiency. The questionnaire was designed with two sections: (1) questions collecting participants' demographic background and personal characteristic features, and (2) questions inquiring participants' general perception of the smart home automation and preference on proactive smart home assistants (Figure 1).

The first section of the questionnaire collected data on participants' general background information and their personal pro-environmental values and beliefs to build their profiles. These information were utilized to classify the participants and predict their general perception towards proactive smart home systems with different levels of autonomy. It has been identified in previous studies that people's demographic background such as age, gender, education level, and number of occupants in the residence can affect their environmental behaviors (Ding et al. 2017) and perceptions towards smart home automation (Singh et al. 2018). As such, in the questionnaire we collected this information as potential influential factors. As half of the energy usage in buildings is from indoor climate conditioning (Chua et al. 2013), previous studies have also investigated the significant impact of occupants' thermal preferences on their decisions in energy-saving behavior (Jazizadeh et al. 2014; Jung and Jazizadeh 2020). In this study we collected participants' indoor thermal preference in cooling seasons with preferred thermostat setpoints, upper limits and lower limits, based on which we computed their thermal preference range and their thermal tolerance type (heat tolerable, neutral, or cold tolerable).

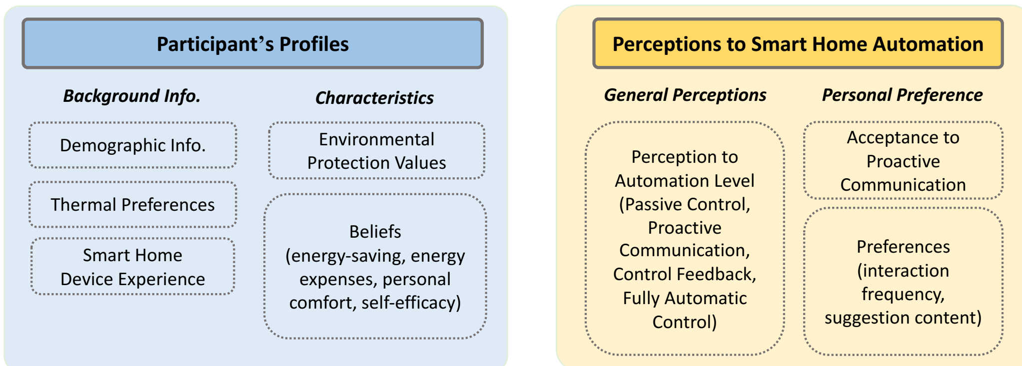

Figure 1. Questionnaire structure and components

Previous studies have also identified that users' previous experience and usage of smart home devices have a close correlation with their future usage and adaptation of the smart home technologies (Sciuto et al. 2018). As such, we also collected data to characterize participants' previous experience with the smart home devices, including the number of smart home devices they have owned, their ownership period, command frequency, frequency of using the smart home interfaces to control appliances, and the number of systems/appliances connected with the smart home devices. Furthermore, the value-belief-norm theory states that one's personal values guide his/her beliefs about human-environment relationship, which further generate his/her personal norms to take actions in energy-saving (Heydarian et al. 2020). As such, we included the values and beliefs in the personal characteristics section and collected participants' values about environmental protection and their beliefs in energy saving, energy expenses, personal comfort and self-efficacy (interest in taking actions for energy-saving).

Apart from the participants' personal profile, we also collected their perceptions of smart home automation and preference on proactive smart home assistants, which was covered in the second section of the questionnaire. First, we collected data on how receptive participants are to smart home systems with smart home assistants at different levels of autonomy: (1) Passive control – only receive commands;(2) proactive communication – give suggestions and wait for occupants' commands; (3) control feedback - automatically adjust appliances and ask for occupants' feedback; (4) fully automatic control – automatically control the system and won't modify until receive occupants' request. In addition, data on participants' acceptance of proactive form of communication from SHAs was collected. In other words, their acceptance of SHAs starting the conversation and giving suggestions was measured. This helped us evaluate whether participants perceive this level of autonomy (taking a more active role) as intrusive. In addition, we further collected data on participants' preference for SHA interactions including frequency of interactions, and the information provided by an SHA, such as cost saving, energy-saving, CO_2 emission reduction, or tips on adaptive behavior.

Data collection and model development. Through an online questionnaire distribution platform, Qualtrics, we collected 58 valid responses as the dataset for further analysis. Data preprocessing were conducted with Likert-scale-type responses transformed into numeric values and all variables in participants' profile section were normalized before being used in training the machine learning models. For the participants' perceptions towards smart home automation and preference on smart home assistants, in most of the cases (unless has specified otherwise in the results section) we labeled them with binary classes: Positive attitude class included the "strongly agree/like" and "somewhat agree/like" responses, and negative attitude class included the "neutral", "somewhat disagree/dislike", and "strongly disagree/dislike" responses.

The pre-processed data contained sixteen variables that represent participants' personal background and characteristics. Using these variables, we developed machine learning models to investigate if we can predict occupant's general acceptance to different levels of smart home autonamy and preference on interaction with proactive smart home assistants. Three widely used classification models were trained and tested in this study: support vector machine (SVM), random forest (RF), and logistic regression (LR). The classification accuracy and F-score were computed as the performance indicators for the trained models. In addition, repeated k-fold cross validation (k = 5, repeated times = 10) were conducted to gain a more accurate estimation of the model performance. Feature importance scores were also computed during the classification, based on which we identified the most important features in the model development.

RESULTS AND FINDINGS

Sample Characteristics. The general sample characteristics of the 58 participants are shown in Table 1, with sixteen variables about participants' profile. Among 58 participants, 35 of them (60%) didn't have any previous experience with smart home devices, and 23 of them (40%) had some experience. For values and beliefs, participants held a generally positive attitude towards environmental protection and energy-saving. Ten indicators of participants' perception towards smart home automation are shown in Table 2. Participants held different attitudes towards smart homes with various levels of autonomy. They were most receptive to the concept of smart homes with proactive communication, while they were least receptive to the concept of smart home systems having full autonomy in control. Among different energy-saving suggestion messages, participants preferred contents with monetary savings, while showing the least interest in suggestions with tips on adaptive behavior.

Table 1. Participants' characteristic features in our sample

Components		Distribution
Demographic Info.	Gender	Males: 31 (53%), Females: 27 (47%)
	Age	21-25: 34 (59%), 26-35: 22 (38%), >35: 2 (3%)
	Education Level	Some College: 1 (2%), Bachelor: 24 (41%), Master: 23 (40%), Doctor: 10 (17%)
	Number of Occupants	Live by oneself: 7 (12%), Live with roommates: 46 (80%), Live with parents: 5 (8%)
Thermal Preference	Thermal Comfort Range	Mean: 6.5°F, Std: 4.26°F
	Thermal Tolerance	Heat tolerable: 23 (40%), Neutral: 15 (26%), Cold tolerable: 20 (34%)
	Number of Devices	None: 35 (60%), At least one: 23 (40%)
	Ownership Period	Zero: 42 (72%), More than one month: 16 (28%)

Components		Distribution
Smart Home Device Experience	Command Frequency	Rarely: 42 (72%), Sometimes: 13 (22%), Often: 3 (6%)
	Control Appliances Frequency	Never: 44 (76%), Sometimes: 3 (6%), Often: 11 (18%)
	Number of Connected System	Zero: 42 (72%), One: 6 (12%), At least two: 10 (16%)
Values and Beliefs	Values (Environmental Protection)	Positive: 52 (90%), Negative: 6 (10%)
	Beliefs (Environmental Protection)	Positive: 53 (91%), Negative: 5 (9%)
	Beliefs (Energy Expenses)	Positive: 51 (88%), Negative: 7 (12%)
	Beliefs (Personal Comfort)	Positive: 33 (57%), Negative: 25 (43%)
	Beliefs (Self-efficacy)	Positive: 45 (76%), Negative: 13 (24%)

Table 2. Participants' response to questions on smart home automation

Sections	Components		Mean	Std.
Perception of Different Autonomy Levels	Passive Control		3.76	0.92
	Proactive Communication		4.10	0.79
	Control Feedback		3.21	1.14
	Fully Automatic Control		2.66	1.29
Personal Preference	Acceptance to Proactive SHAs		3.57	0.92
	Interaction Frequency		2.31	0.98
	Feedback Content	Cost Saving	4.33	0.69
		Energy Saving	4.03	0.72
		Emission Reduction	3.76	0.90
		Adaptive Behavior Tips	3.62	1.21

Machine Learning Models. Three machine learning models were trained to predict participants' preference on smart home with different levels of autonomy (Table 3), preference of interaction form (Table 4), and preference on the content of suggestions from SHA (Table 5).

For preference on smart home levels of autonomy, we conducted both a three-class and a two-class classification. In three-class classification, participants were grouped into (1) Prefer passive control, (2) Prefer proactive communication, and (3) Prefer automatic control, while in the two-class classification they were grouped into (1) Prefer passive control, and (2) Open to higher level of autonomy. As shown in Table 3, models did not perform well in three-class classification. However, the two-class classifiers were relatively successful with a good performance, among which random forest outperformed the rest with an accuracy of 0.82 and a F-score of 0.77.

We also computed the feature importance scores using the random forest model and identified five features with the highest scores (Table 3). It can be seen that participants' self-efficacy belief and environmental protection belief play an important role in distinguishing between the two groups. Self-efficacy belief indicates participants' interest in taking actions in energy-saving if proper guidance is provided, and it is shown by the feature importance score that this feature has considerable effect on participants' acceptability to smart homes with higher level of autonomy. Also, participants' environmental protection value showed high explanatory power in the model.

Table 3. Automation level acceptance - performance of models and important features

Classes	Performance	SVM	RF	LR	Top Five Important Features
Three-Class Classification	Accuracy	0.47	0.45	0.48	-
	F-Score	0.37	0.41	0.38	
Two-Class Classification	Accuracy	0.82	0.82	0.81	Belief (Self-efficacy), Belief (Environmental Protection), Thermal comfort range, Age, Value (Environmental Protection)
	F-Score	0.75	0.77	0.74	

In classification of the interaction form preference, we tried to classify participants based on being generally receptive to proactive smart home assistants giving energy-saving suggestions and their preferred interaction frequency (Table 4). The results showed that the models had a good performance (with 0.73 of accuracy and 0.71 of F-score in random forest model) in predicting participants' being receptive to the proactive SHAs giving energy-saving suggestions. However, the classification did not work well in predicting the preferred frequency of interactions (with the highest accuracy score of 0.58 from the random forest model). As such, we only identified the top five important features in the acceptance prediction model. Similar to the influential factors in the previous case, participants' self-efficacy and energy expenses beliefs, and their environmental protection value ranked high in the feature importance scores.

Table 4. Interaction preference - performance of models and important features

Interaction Preference	Measure	SVM	RF	LR	Top Five Important Features
Acceptance to Proactive SHAs with Energy Suggestions	Accuracy	0.67	0.73	0.69	Belief (Self-efficacy), Belief (Energy Expenses), Values (Environmental Protection), Age, Belief (Environmental Protection)
	F-Score	0.66	0.71	0.66	
Interaction Frequency	Accuracy	0.48	0.58	0.49	-
	F-Score	0.47	0.58	0.48	

In the case of feedback contents, we treated each type of feedback as one classification problem. All three models can classify participants' preference on suggestions that have cost saving contents with high accuracy (0.91) and F-scores (0.88). The most important feature in identifying the participants that prefer cost saving contents is energy expenses' belief. This observation is expected because since if people believe in the effect of their daily behaviors on energy cost saving, they would be more interested in the suggestions with the actual number of savings on their energy bills. On the other hand, for suggestions with information on energy saving and CO_2 emission reduction, participants' environmental protection value plays a more important role. Participants with environmental-friendly values would be more sensitive to their daily carbon footprint and thus show more interest in suggestions with contents about positive environmental impact.

As discussed above, the significant impact of participants' pro-environmental values and beliefs on their perception towards smart home automation aligns with the value-belief-norm theory in previous studies. As such, smart home assistants can proactively collect occupants' profiles about their environment and energy-related values or beliefs to develop predictive models that enable them with tailored communication with different users.

Table 5. Suggestion content preference - performance of models and important features

Suggestion Contents	Measure	SVM	RF	LR	Top Five Important Features
Cost Saving	Accuracy	0.91	0.91	0.91	Belief (Energy Expenses), Belief (Self-efficacy), Number of Occupants, Belief (Environmental Protection), Thermal Comfort Range
	F-Score	0.88	0.87	0.88	
Energy Saving	Accuracy	0.79	0.78	0.79	Values (Environmental Protection), Belief (Self-efficacy), Education, Belief (Energy Expense), Age
	F-Score	0.70	0.72	0.71	
CO2 Emission Saving	Accuracy	0.67	0.77	0.67	Values (Environmental Protection), Belief (Energy Expenses), Belief (Self-efficacy), Thermal Comfort Range, Education
	F-Score	0.65	0.73	0.64	
Adaptive Behavior Tips	Accuracy	0.57	0.62	0.59	-
	F-Score	0.52	0.62	0.52	

CONCLUSION

In this study, through an online experiment in a campus community, and by utilizing machine learning techniques, we identified important personal features with an impact on users' preferences on proactive mode of communication from smart home virtual assistants, tailored content of communication, and different levels of autonomy. Participants, in this study, were more receptive to smart home systems with proactive communication compared to passive control that is initiated by users. Nevertheless, they were less willing to let the smart home have fully automatic control over their house appliances.

Prediction models indicated that the smart home assistants can learn from participants' profiles and predict their preferences on the autonomy mode with a good performance. In predicting participants' attitudes (positive versus negative) towards proactive smart home assistants with energy-saving suggestions, the random forest model performed better than other models with an accuracy of 0.73 and F-score of 0.71. The value-belief-norm theory was identified to be very applicable and reliable in designing the learning frameworks for energy-related interactions initiated by smart home assistants. To this end, participants' pro-environmental values and energy-saving-related beliefs showed higher predictive power compared to other variables. Participants' beliefs in self-efficacy and energy expenses, and environmental protection values were important features in identifying occupants' general attitude and preferences of feedback contents.

AKNOWLEDGEMNT

This material is based upon work supported by Virginia's 4-VA Collaborative Research Grant and the National Science Foundation under grant #1663513. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation and 4-VA program.

REFERENCES

Alaa, M., Zaidan, A. A., Zaidan, B. B., Talal, M., and Kiah, M. L. M. (2017). "A review of smart home applications based on Internet of Things." *Journal of Network and Computer Applications*, 97, 48-65.

Chua, K. J., Chou, S. K., Yang, W., and Yan, J. (2013). "Achieving better energy-efficient air conditioning–a review of technologies and strategies." *Applied Energy*, 104, 87-104.

Ding, Z., Wang, G., Liu, Z., and Long, R. (2017). "Research on differences in the factors influencing the energy-saving behavior of urban and rural residents in China–A case study of Jiangsu Province." *Energy Policy*, 100, 252-259.

Georgiev, A., and Schlögl, S. (2018). "Smart home technology: An exploration of end user perceptions." *Innovative Lösungen für eine alternde Gesellschaft: Konferenzbeiträge der SMARTER LIVES*, 18(20.02).

Heydarian, A., McIlvennie, C., Arpan, L., Yousefi, S., Syndicus, M., Schweiker, M., Jazizadeh, F., Rissetto, R., Pisello, A. L., and Piselli, C. (2020). "What drives our behaviors in buildings? A review on occupant interactions with building systems from the lens of behavioral theories." *Building and Environment*, 106928.

Jazizadeh, F., Ghahramani, A., Becerik-Gerber, B., Kichkaylo, T., and Orosz, M. (2014). "Human-building interaction framework for personalized thermal comfort-driven systems in office buildings." *Journal of Computing in Civil Engineering*, 28(1), 2-16.

Jivani, F. D., Malvankar, M., and Shankarmani, R. "A Voice Controlled Smart Home Solution With a Centralized Management Framework Implemented Using AI and NLP." *Proc., 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, IEEE, 1-5.

Jung, W., and Jazizadeh, F. (2020). "Energy saving potentials of integrating personal thermal comfort models for control of building systems: Comprehensive quantification through combinatorial consideration of influential parameters." *Applied Energy*, 268, 114882.

Mennicken, S., Zihler, O., Juldaschewa, F., Molnar, V., Aggeler, D., and Huang, E. M. "" It's like living with a friendly stranger" perceptions of personality traits in a smart home." *Proc., Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 120-131.

Miksik, O., Munasinghe, I., Asensio-Cubero, J., Bethi, S. R., Huang, S., Zylfo, S., Liu, X., Nica, T., Mitrocsak, A., and Mezza, S. (2020). "Building proactive voice assistants: When and how (not) to interact." *arXiv preprint arXiv:2005.01322*.

Sciuto, A., Saini, A., Forlizzi, J., and Hong, J. I. "" Hey Alexa, What's Up?" A Mixed-Methods Studies of In-Home Conversational Agent Usage." *Proc., Proceedings of the 2018 Designing Interactive Systems Conference*, 857-868.

Singh, D., Psychoula, I., Kropf, J., Hanke, S., and Holzinger, A. "Users' perceptions and attitudes towards smart home technologies." *Proc., International Conference on Smart Homes and Health Telematics*, Springer, 203-214.

Tabassum, M., Kosinski, T., and Lipford, H. R. "" I don't own the data": End User Perceptions of Smart Home Device Data Practices and Risks." *Proc., Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*.