Building and Environment 170 (2020) 106603

Contents lists available at ScienceDirect

Building and Environment

journal homepage: http://www.elsevier.com/locate/buildenv

An assessment of opinions and perceptions of smart thermostats using aspect-based sentiment analysis of online reviews

Diba Malekpour Koupaei [a], Taikgun Song [b], Kristen S. Cetin [c,*], Jongho Im [d]

[a] Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA, USA
[b] College of Business, University of Florida, Gainsville, FL, USA
[c] Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA
[d] Department of Applied Statistics, Yonsei University, Seoul, South Korea

ARTICLE INFO

Keywords:
Energy savings
Smart thermostats
Smart home technology
Home energy management
Opinion mining

ABSTRACT

Smart thermostats have been on the market for nearly a decade, with an estimated adoption rate of 7% in 2018. With many regions of the U.S. having a heating and/or cooling system in nearly 100% of households, there is significant opportunity for further adoption, which can help support energy savings and building-grid interactions. However, more insight is needed to provide a better understanding of their utilization and user opinions. In this study, online reviews are used to evaluate users' perceptions and attitudes towards smart thermostats. 26,372 product reviews were collected for five commercially-available smart thermostats and were analyzed with a confirmatory aspect-based opinion mining technique. An analysis of this dataset shows that the characteristics of the current user population show substantial differences compared to the more widely studied early adopters. When comparing the most commonly discussed topics, users generally do not discuss the energy and cost savings related features of their devices in comparison to other topics such as control, ease of use, and installation. In addition, comfort is discussed nearly twice as much as energy efficiency. The results of this work can help product manufacturers and utility providers to push towards more widespread adoption and efficient use.

1. Introduction

By 2040, it is projected that worldwide energy consumption will increase by 28% [1]. In the U.S. alone, energy use is predicted to grow at an even greater rate of over 0.4% per year over the next three decades [2]. Such projections put further emphasis on the need for achieving reduction goals in energy demand. The building sector represents a significant opportunity for such reduction measures, as it is a major contributor to the current energy and electricity demands [3]. In this context, reduction measures include but are not limited to efficiency and operational improvements to buildings' energy-consuming systems [4]. Residential buildings in particular offer a unique opportunity for improvements as they represent over half of the U.S. building stock's energy consumption [3]. Moreover, their energy consuming systems are often neither serviced on a regular schedule to check for inefficiencies, nor operated with the sophisticated technologies that are more commonly used in many commercial buildings [5]. Consequently, recent trends and interest in connected devices and technologies used to operate residential buildings more efficiently can help to significantly improve the energy performance of these buildings [6].

In the past several years, connectivity of people and more recently buildings through the use of connected devices, also called the Internet of Things (IoT), has increased dramatically. Currently in the U.S., it is estimated that more than 80% of households have at least one internet connected device in their home, including 76.5% with at least one smartphone, 77.4% with at least one computer, and 57.8% with at least one tablet [7]. In the context of residential buildings, there has been significant increase in the number of technologies commercially available to make homes more connected and intelligently operated. These "smart home technologies", range significantly in type and functionality, and while the adoption of these devices remains limited to a minority of today's housing units, some technologies have had relatively larger success and are thus more ubiquitous. Overall, it is estimated that approximately 33.2% of households in the U.S. currently have at least one smart home device and market trends predict significant growth in the number of smart home devices in use in the next 5–10 years [8].

* Corresponding author.
E-mail address: cetinkri@msu.edu (K.S. Cetin).

https://doi.org/10.1016/j.buildenv.2019.106603
Received 13 September 2019; Received in revised form 9 December 2019; Accepted 14 December 2019
Available online 17 December 2019

According to some estimates, by 2023 the number of homes equipped with at least one smart device is predicted to increase to over 53% [8].

These smart home devices focus on meeting several needs and goals, such as control, comfort, lighting, security, home entertainment, energy management, and appliance management and control. In the U.S., energy management devices, which include smart and internet-connected thermostats and sensors that detect and/or adjust HVAC (heating, ventilation and air-conditioning) systems and other end-uses, are predicted to be in approximately 21% of homes equipped with a smart device by 2023 [8]. Among other smart home devices, control- and connectivity-related devices are predicted to have the highest penetration rate (51%), followed by home entertainment (36%) and then comfort and lighting devices (33%) [8]. While the adoption of energy management devices does not represent the largest share of current consumer interests among smart home technologies, their impact on energy efficiency still has the potential to be significant. This is due, at least in part, to energy management devices' target area for improvement, which is the HVAC system, being among the largest consumers of energy in residential buildings [9]. Moreover, the HVAC system is often the single largest contributor to electricity demand during peak demand hours, particularly in extreme heat and to a lesser extent, extreme cold conditions [10,11]. Secondly, in the case of homes undergoing retrofits, smart thermostats are the most common smart device installed as part of a renovation or upgrade [12]. Given that existing buildings make up a large portion of the U.S. building stock and tend to be less energy efficient compared to new construction, any improvement in their energy performance can provide an opportunity for energy demand reduction [13].

Accordingly, the focus of this study is on the adoption of smart thermostats (also referred to as "intelligent thermostats") as one type of energy management device that is expected to be more commonly used in the near future, and can also help to significantly reduce energy demands in residential buildings. Generally speaking, a "smart thermostat" has the capability to remotely control the setpoints as well as to adjust setpoint schedules of the HVAC system. In some cases, it can also monitor occupant behavior and learn from this behavior for improved HVAC operational efficiency and occupant comfort. Some devices can also connect to the local utility company and respond to utility signals to adjust HVAC operations to support peak load reductions [14]. The performance of such devices has been the main focus of a number of recent studies, the majority of which indicate significant impacts from the use of these devices in terms of energy use and/or demand reduction (e.g. Refs. [15,16]). However, the current challenge with translating smart thermostats' energy-saving potential to real results lies in the relatively low penetration rate of these devices [17]. While it is estimated that 80–100% of U.S. residential buildings are equipped with thermostats, the proportional share of smart thermostats is fairly low [18].

Several recent studies have attempted to identify barriers to the adoption of energy efficient technologies. One common theme in their findings is that they cite the overall high capital costs of these technologies as one of the main barriers to their adoption [19,20]. However, some studies also argue that this specific barrier can be overcome by demonstrating the non-energy related benefits of such technologies [21,22]. For instance, Im et al. (2017) demonstrated that the installation of energy efficient technologies has a positive impact on housing and rental prices [23]. Communicating such findings can help potential users make the decision to purchase smart technologies, especially if their energy-related savings are not sufficient to justify the initial investment. Moreover, recent studies indicate that there are non-monetary influential factors involved in people's decision-making processes that can outweigh financial considerations [24]. These influential factors, if identified and quantified, can also help to increase the adoption of smart technologies and thus reduce energy consumption. For smart thermostats, to the best of our knowledge, the only recent publications that have studied the influential factors involved in the decision to purchase and/or use smart thermostats, or the relative considerations of energy and non-energy benefits have focused on early adopters of these technologies. In these studies, Yang et al. (2012; 2013) used phone interviews and diaries with a limited user population of one specific smart thermostat as the basis for their analysis [16,25]. Therefore, given that smart thermostats have been commercially available for nearly a decade, this effort works towards evaluation of a broader set of users' perception and attitudes toward these devices.

The objective of this study is to assess the motivation and sentiment of current buyers and users of smart thermostats through analyzing reviews on the online retail website *Amazon.com*. Five different popular smart thermostats available through this site were chosen. Then a confirmatory aspect-based opinion mining technique, which divides online reviews into aspects and then evaluates reviewers' sentiments associated with each of these aspects, was used to study the 26,372 customer reviews. 2196 reviews on four programmable but non-smart thermostats (hereafter referred to as non-smart thermostats) were also extracted and compared to the smart thermostat results. The use of an online review database has a number of advantages over more conventional survey methods used previously: (a) online reviews are free and publicly accessible, (b) the sample size is larger than most laboratory or field test populations are able to achieve; (c) there is no questionnaire involved in online reviews (that may impact users' responses) thus reviewers discuss what they find relevant to share about a device which may help to capture the reality of user perception towards a product. The results of this study provide insight into the relative interests and cited benefits (energy and/or non-energy related) of such smart thermostats. These results also work toward determining the motivational factors for the purchase of these devices. The latter can be used to inform targeted efforts to further encourage adoption of more energy efficient technologies, and hence overall efficiency of the residential building stock.

2. Methodology

Opinion mining or sentiment analysis refers to a set of techniques that use opinion related data to extract beneficial information. Related literature offers two main approaches, namely aspect-based and non-aspect-based opinion mining [26]. In this study, an aspect-based opinion mining technique is used that divides online reviews (input data) into aspects, also called features or subtopics, and then evaluates reviewers' opinions of each of these aspects. These aspects usually correspond to the selected products' important features and therefore enable detailed analysis of the reviewers' opinion of the studied products.

Of the three types of text summarization methods, including aspect detection, sentiment analysis, and joint aspect detection and sentiment analysis [27], joint aspect detection and sentiment analysis methods are beneficial for data where topics of interest are not pre-defined. However, it is challenging to mine opinions from reviews that are in natural language, particularly if product aspects are not directly presented within the reviews [28]. Moreover, sentences may contain multiple aspects and the reviewers' opinion on each of these aspects may be unclear. To overcome such challenges, a confirmatory aspect-based opinion mining technique is used, proposed by Im et al. (2019), that employs the concept of Confirmatory Factor Analysis (CFA) [29]. In this technique, in contract to all previously proposed methods including both machine-learning based algorithms and lexicon-based approaches, word (or phrase) seeds as the potential aspects are given in advance, which enabled the use of domain knowledge to focus on features of interest. The following subsections describe the data collection and processing procedure in further detail.

2.1. Data collection

In order to analyze homeowners' interests and opinions of smart thermostats, review texts and ratings were extracted from *Amazon.com* using the web scraping R package *rvest* in September 2017 and February 2018 [30,31]. In total, 26,372 product reviews were collected for the following five commercially available smart thermostats: Sensi Smart Thermostat 524547878, Nest Learning Thermostat T3007ES (3rd Gen), Ecobee3 EB-STATe3-O2 (2nd Gen), Honeywell Wi-Fi 7-Day Programmable Thermostat RTH9580WF 1005/W1, and Honeywell Wi-Fi 7-Day Programmable Thermostat RTH6580WF1001/W1. The criteria for choosing these five devices was that their manufactures were identified as dominant market players (based on [14]) and the devices themselves had high overall ratings (at least 4 out of 5 stars) on *Amazon.com*. As mentioned, smart thermostats generally have been defined as having the capabilities and features to remotely control the thermostat and in some cases monitor occupant behavior and learn from this behavior for improved efficiency and comfort [14]. All of the selected thermostats have similar features that fit within these guidelines, including Wi-Fi connectivity, programmable 7-day flexible scheduling, a digital display, and a cell phone application to enable remote control, checking of system operations, and facilitating programming capabilities. In addition, they are all designed to be used for control of a single zone residential style HVAC system, including a heat pump or air conditioner/furnace combination which is commonly used in residential buildings in the U.S [18]. Some of the selected smart thermostats had additional features including occupancy sensing and learning algorithms, geofencing, alerts for extreme indoor temperatures, touch screens, and the ability to connect to remote temperature sensors placed in different locations within a home. These minor variations among features and capabilities were unavoidable. Generally, regardless of these variations, all of the selected thermostats can be classified as smart thermostats and have an average rating of at least 4 out of a 5-point scale in the online retailer's website. In addition, a sufficient number of online reviews per device was also a requirement to provide a strong dataset for analysis. Table 1 presents a summary of the collected dataset for the five selected smart thermostats including the number of reviews acquired, their average overall ratings, and the associate standard errors.

Data was also collected in a similar fashion for four programmable (non-smart) thermostats to enable comparison between the two types of thermostats. These non-smart thermostats were chosen based on the total number of reviews, where thermostats with the most reviews were prioritized over those with less reviews and in total, 2196 product reviews were collected in this group, including data for the following thermostats: Honeywell RTH2300B1012, Honeywell RTH221B1021, Lux TX500U, and Hunter 44157. A comparison of reported prices in Tables 1 and 2 indicates that among the thermostats studied, the cost of the smart thermostats ranged from $85 to $250. While this is a broad range of prices, they are all at a higher price point than a non-programmable thermostat (generally ranged from $10 to $25) and a non-smart programmable thermostat (generally ranged from $15 to $40). Other important differences between the two datasets are the lower number of reviews available for the non-smart thermostats available during the time of data collection as well as slightly lower overall ratings when compared to the smart ones. Table 2 provides a summary of the collected dataset for the selected non-smart thermostats.

Table 2
Summary of the acquired review dataset for non-smart thermostats.

Product Name	Price (US$ 2018)	Number of Reviews	Average Rating (out of 5)	Standard Deviation
Non-Smart Thermostat A	$36.99	624	3.94	1.40
Non-Smart Thermostat B	$24.99	1,142	4.15	1.31
Non-Smart Thermostat C	$35.00	79	3.00	1.79
Non-Smart Thermostat D	$17.91	351	3.90	1.45

2.2. Data cleaning and processing procedures

The acquired dataset was then refined to improve the analysis. This cleaning and processing procedure includes the following four steps:

1. Removal of reviews written in non-English languages;
2. Removal of *stop words*, which often refers to non-meaningful common words such as 'the', 'a', and 'my', and non-meaningful punctuation marks;
3. Removal of irrelevant punctuation marks. This includes punctuation marks such as '!' and '?' used multiple times in a consecutive manner (as a form of exaggeration), punctuation marks used to construct a pictorial representation of an image (also known as emoticons), and apostrophes used as a form of abbreviation;
4. Application of a *stemming* procedure. This procedure converts the words to their morphological word form.

The first step was deemed necessary since some users had provided reviews in languages other than English. Since the non-English words cannot be interpreted using the same methodologies followed for English-based text analysis, these reviews were removed from the overall database of reviews. Otherwise, the non-English reviews would have interrupted the technique's ability to obtain consistent results. In the second step, stop words were also removed from the remaining reviews, since their inclusion in a textual analysis provides unnecessary information and clutter in the text dataset. This was realized using the *tm* package in R [31,32]. For the irrelevant punctuation marks (Step 3), while these punctuation marks are a possible means to express sentiment, they are not used in the main data mining technique used in this work and could be studied further in future efforts. In addition, words with an apostrophe to indicate possession, omission, and pluralization were converted to words without an apostrophe. For example, "didn't" was changed to "did not".

Following the cleaning procedure, a fourth and final step called stemming was applied to the data to reduce the number of unique words in the dataset. For example, words such as "stems", "stemming", and "stemmed" would all be expressed with the same root word "stem" instead of their original forms. This is necessary because the opinion-mining algorithm used herein requires natural language process (NLP) and tagging of part-of-speech (POS) associated with each word.

Table 1
Summary of the online reviews dataset for smart thermostats.

Product Name	Price (US$ 2018)	Number of Reviews	Average Rating (out of 5)	Standard Error	Unique Features
Smart Thermostat A	$142	3,153	4.21	1.35	Geofencing
Smart Thermostat B	$241	14,786	4.74	0.82	Learning Algorithm, Touch Screen
Smart Thermostat C	$250	3,965	4.54	1.06	Learning Algorithm, Touch Screen, Remote Sensors
Smart Thermostat D	$159	2,870	4.17	1.36	Touch Screen
Smart Thermostat E	$85	1,598	4.17	1.37	Learning Algorithm

2.3. Application of the selected opinion mining technique

Earlier, it was mentioned that the selected method for this study is a confirmatory aspect-based opinion mining technique that requires pre-defined aspect categories. Here, each review is fitted into one of the following eight aspect categories: "*feelings*" (X_1), "*installation*" (X_2), "*easy/hard*" (X_3), "*control/occupancy*" (X_4), "*connectivity/application*" (X_5), "*energy*" (X_6), "*costs*" (X_7), and "*comfort/HVAC*" (X_8). These categories of bi-terms were chosen based on themes of specific features or elements of the thermostats in the dataset, developed based on reviewing the dataset and combining the most commonly utilized bi-terms associated with similar features together (Fig. 1). These same categories were used for analysis of both the smart and non-smart thermostats so as to enable a parallel comparison of the two datasets.

Following the establishment of these topic (aspect) categories, for each one of the collected reviews, it was first determined if that specific review included words associated with each of these topic categories. If there were no words associated with a certain aspect category, then it was denoted as 'N/A' (Table 3). Then, for all the aspects not tagged 'N/A', we determined if the sentiment of the review towards that aspects was positive (P), negative (N), or neutral (U). Table 3 illustrates an example of this process and its results for three different reviews in the developed database.

For example, in the first review (Review ID = 1), the text associated with the aspect domains X_1 to X_3 are all positive overall, while the X_5 related text is negative. There are no sentences or clauses in this review which are directly or implicitly linked to the domain X_4 and all other topics' texts are neutral (related to the topic domain, but not positive or negative). The following subsections discuss this process in more detail.

2.4. Detailed description of the sentiment analysis procedure

The confirmatory aspect-based opinion mining technique used in this study utilizes NLP to extract users' opinions. While in conventional aspect-based opinion mining the product aspects are automatically investigated based on the text itself, the proposed approach allows pre-determined product aspects (such as the eight topic categories and related key words often referred to as seeds) to be provided in advance [33]. The proposed algorithm consists of the following six substeps as shown in Fig. 2: disintegrating, summarizing, straining, bagging, upcycling, and scoring.

First, customer reviews are separated into a set of clauses (disintegrating). Second, the disintegrated clauses are summarized into a set of bi-terms (summarizing) using a part-of-speech tagger. Third, bi-terms are strained based on their frequencies (straining). Fourth, the strained bi-terms are manually matched to the pre-determined topic categories using external information or domain knowledge (bagging). Fifth, using the bag-of-topics obtained in the fourth step, unstrained bi-terms are also matched to the given topic domains and therefore bags-of-topics are expanded (upcycling). Finally, sentiment scores are assigned to the clustered bi-terms (scoring).

Table 3
Examples of the aspect-based sentiment analysis.

Review ID	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	P	P	P	N/A	N	U	U	U
2	N/A	N/A	P	P	U	P	U	P
3	U	N	U	P	P	P	P	P

Note: (P) positive; (N) negative; (U) neutral; and (N/A) not applicable.

To better demonstrate the functionality of these six steps, the following is an example of one of the thermostat reviews used to discuss the complete process.

> "Easy to install and very happy with the energy savings. The daily schedule is perfect for our family needs and I have seen considerable monthly savings. Remote management through the android app is simple and easy to use. I have had no issues with the device or the app."

2.4.1. Disintegrating

Because sentences may include several aspects, it is more effective to apply this algorithm at the clause level. Therefore, in this step, each review is separated into a set of clauses and each clause is assigned one of the eight aspect categories. Clause segmentation or identification is similar to classical sentence segmentation that separates a review into a set of sentences based on punctuation marks. The difference is that clause segmentation requires additional clause boundaries around conjunction word such as "and". Accordingly, when the words before or after a conjunction in a sentence have a different POS, the sentence is considered to consist of two clauses. In other words, detecting clause boundaries is implemented by identifying the POS of the words that are located near conjunction words.

2.4.2. Summarizing

The challenge with clauses is that they include too many words and mining more than two words at any instance is challenging and inefficient. Therefore, to enhance computational efficiency, each clause is then summarized into a set of bi-terms which consist of (a) an aspect word and (b) an opinion word. For example, an input clause, "*I am extremely happy with the thermostat*", can be summarized as (happy, thermostat), where 'thermostat' represents the aspect word and 'happy' is the opinion word. According to rules introduced by Im et al. (2019), the following five major word combinations are used to summarize each review: (Noun, Adjective), (Noun, Verb), (Noun, Adverb), (Verb, Adjective), and (Verb, Adverb) [29]. Table 4 demonstrates the results of this step for the selected example review.

2.4.3. Straining

Ultimately, the review database resulted in a total number of 105,648 unique bi-terms. These bi-terms were then separated based on

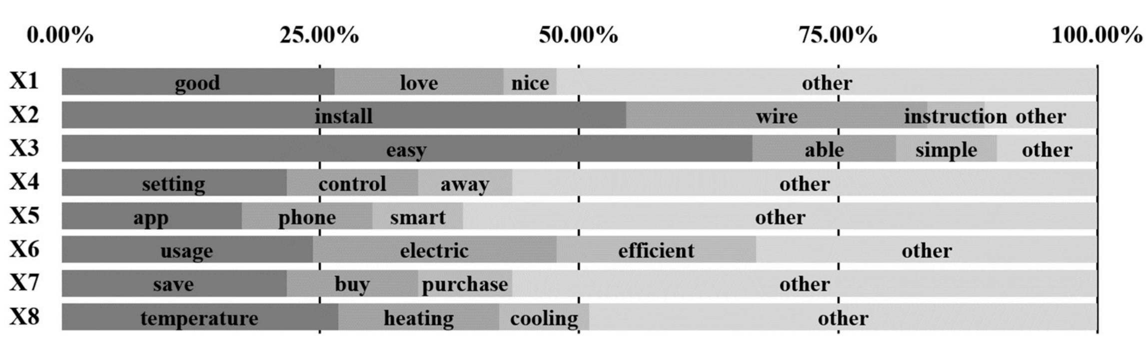

Fig. 1. Top 3 keywords associated with each of the topic categories. (Note: the topic categories are as following: "feelings" (X1), "installation" (X2), "easy/hard" (X3), "control/occupancy" (X4), "connectivity/application" (X5), "energy" (X6), "costs" (X7), and "comfort/HVAC" (X8)).

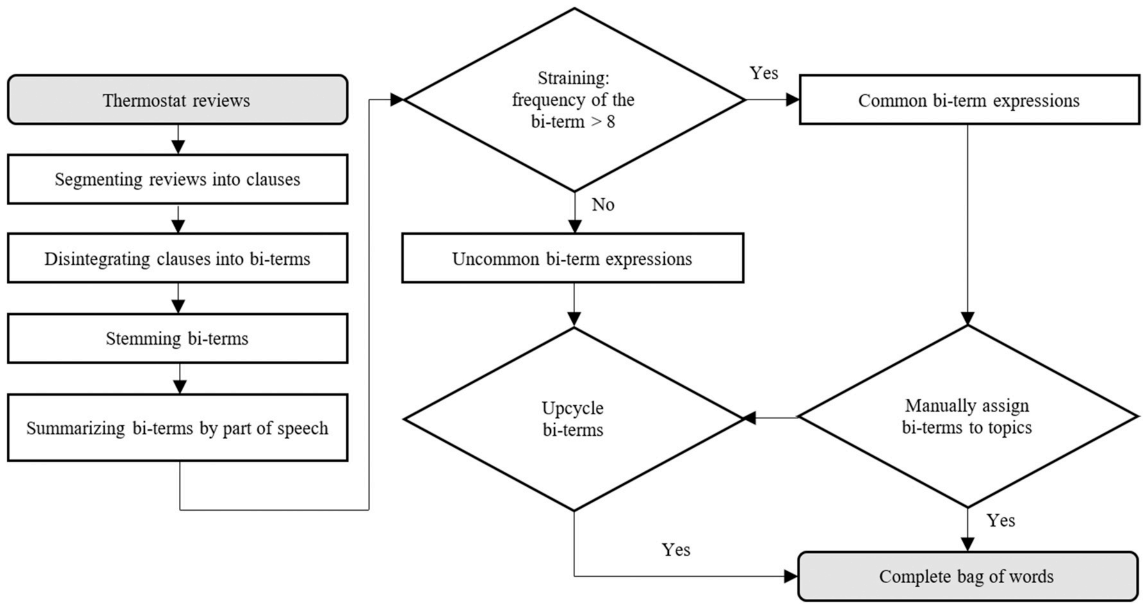

Fig. 2. Confirmatory aspect-based opinion mining technique and its substeps.

Table 4
Bi-term results for the example review.

Aspect Word	Part of Speech	Opinion Word	Part of Speech
instal	verb	easi	adjective
use	verb	easi	adjective
thermostat	noun	old	adjective
work	verb	great	adjective
sensor	noun	remot	adjective
work	verb	well	adverb
money	noun	save	verb
instal	noun	easi	adjective
thermostat	noun	smart	adjective
product	noun	great	adjective

their frequency of appearance, such that the most common bi-terms (i.e. bi-terms that were used more than eight times throughout the database) were labeled as 'common' and all others were labeled as 'uncommon'. This procedure is named 'straining' because it effectively strains the bi-terms to the most frequently used bi-terms across reviews. This step is commonly known as the labeling step for supervised learning.

2.4.4. *Bagging*

In the bagging step, each common bi-term is manually assigned into one (or more than one) of the eight pre-defined aspect categories. Table 5 provides a list of the ten most common bi-terms across all smart thermostat reviews in the dataset. Thus far the straining and bagging steps together have constructed an initial bag-of-bi-terms that will be used as a learning model on the bi-terms and match them with the fixed eight aspect categories.

Table 5
Top 10 most common bi-terms in our dataset and their assigned topic categories.

Count	Bi-terms	Frequency (%)	Count	Aspect(s)
1	instal-easi	2.42	13,506	Installation Easy/Hard
2	use-easi	1.00	5,534	Easy/Hard
3	work-great	0.90	5,026	Feelings
4	thermostat-love	0.85	4,745	Feelings
5	product-great	0.80	4,474	Feelings
6	product-excel	0.70	3,880	Feelings
7	money-save	0.53	2,934	Cost; Energy
8	set-easi	0.48	2,664	Easy/Hard
9	setup-easi	0.44	2,472	Easy/Hard
10	product-love	0.34	1,915	Feelings

2.4.5. *Upcycling*

Other than the common bi-terms, a unique and uncommon bi-term can also be informative with respect to the pre-defined aspect categories. Therefore, to find such bi-terms among uncommon bi-terms and include them in this analysis, a substep called upcycling is completed. As an example, the important bi-term "instruct-thorough" appeared only seven times in the original review database. During the upcycling step this uncommon bi-term was matched with the following three common bi-terms: "instruct-detail", "instruct-simpl", and "instruct-complet". These three common bi-terms were assigned to "Installation" and "Easy/Hard" aspect categories. Therefore, "instruct-thorough" bi-term was also assigned to both "Installation" and "Easy/Hard" aspect categories. Table 6 shows the results of this step for our selected example review.

This procedure is conducted autonomously using external information through an online dictionary. These information sources serve as a reference that determines if two different words convey the same thing. This can be completed using a binary value (0 or 1) for each comparison made between a selected uncommon and common bi-term and then the scores are merged within the bag-of-aspects. After this comparison process, the most suitable aspect category is selected for the given uncommon bi-term based on comparing the averaged comparison scores for all aspect categories.

Table 6
Upcycling results for the selected example review.

Aspect Word	Part of Speech	Opinion Word	Part of Speech	Aspect
instal	noun	fast	adjective	Installation
instruct	noun	thorough	adjective	Installation
thermostat	noun	capabl	adjective	Feelings
way	noun	easier	adjective	Easy/Hard
one	noun	better	adjective	Feelings
temp	noun	consist	adjective	Comfort/HVAC
inform	noun	detail	adjective	Installation
sensor	noun	amaz	adjective	Feelings
sensor	noun	fantast	adjective	Feelings
schedul	noun	intuit	adjective	Control/ Occupancy

2.4.6. Scoring

The final step in this procedure is to assign a sentiment score to each bi-term set. This score is calculated using sentiment lexicons. In this study, the WordNet lexical database and the *tidytext* package included in R is used as the sentiment lexicon [31,34,35]. While it is preferred and beneficial to use domain-specific dictionaries, one does not currently exist for this context, thus a general sentiment dictionary was used. A field specific sentiment dictionary would be beneficial to consider for future analyses, and is the subject of future work. Sample sentiment scores for the example review are reported in Table 7.

3. Results

This section presents the results by applying the opinion mining technique to the two databases of acquired online reviews for smart and non-smart thermostats. Examples from the data are added to the text for improved clarity. Each referenced review is designated using "STR" for a smart thermostat review(er) or "PTR" for a non-smart, programmable thermostat review(er), and a review number. For example, STR #127 is the 127th smart thermostat review(er). No edits were made to the reviewer comments, even when typos were present. First, the smart thermostat reviews are discussed, followed by non-smart thermostats', with the goal of better understanding the motivational factors and sentiment of thermostat purchasers.

3.1. Smart thermostat reviews

Using the proposed opinion mining technique, nearly 75% of all bi-terms collected from smart thermostat reviews were categorized into the following eight topic domains: "*Feelings*", "*Control/Occupancy*", "*Easy/Hard*", "*Connectivity/Application*", "*Installation*", "*Comfort/HVAC*", "*Costs*", and "*Energy*". Fig. 3 shows the frequency of these bi-terms in the smart thermostat reviews database per topic domain. The data also represents the sentiment score assigned to the bi-terms (positive, negative, or neutral).

Overall, the *control/occupancy* topic category was one of the most common topics, at 22.2% of total categorized bi-terms. This is not surprising given that one of the main differentiating features of smart thermostats over the previous generation of programmable or manual thermostats is their extensive control and/or automation features. Therefore, it is understandable that users were enthusiastic to explore and comment on this capability. The majority of reviews associated with this category were considered neutral (64.8%), some were positive (26.3%), and a significantly smaller amount were negative (8.9%). This seems to indicate that while the reviewers demonstrate interest in their device's control features, their sentiment towards this capability is not yet well defined, but is generally not negative. Previous research on smart home technologies suggests that control features have both benefits and criticisms [36]. Many reviewers appreciate the thermostat's controls and automation, such as STR #9203 who states "*… probably the most important reason for owning one is the money it saves me by automatically controlling the temperature in the house.*" However, others indicate discomfort with the controls [36], such as STR #2269 who expresses his/her concerns as: "*I still cling to the notion that I'm smarter than any thermostat, so I keep my [device name] in "hold" mode and adjust it manually (though often remotely) …*" These demonstrate the interest in controls and automation, but also the somewhat conflicting opinions around this topic. The optimal path suggested by others, and supported by these findings, may be to provide choices which allow for a balance between desired user control and automated features [37]. However, given the overall high level of discussion on these features relative to other topics, and the generally positive or neutral sentiment associated with this discussion, these features appear to be important and a motivation, at least for some, for the purchase and/or use of smart thermostats.

Table 7
Scoring results for the selected example review.

Aspect Word	Part of Speech	Opinion Word	Part of Speech	Assigned Topic	Sentiment Score
bundle	noun	smarter	adjective	Feelings	positive
bundle	noun	smarter	adjective	Control/ Occupancy	positive
heat	noun	electric	adjective	Comfort/ HVAC	neutral
thermostat	noun	purchased	verb	Costs	neutral
product	noun	excellent	adjective	Feelings	positive
access	noun	remote	adjective	Connectivity/ application	neutral
access	noun	remote	adjective	Control/ Occupancy	neutral
features	noun	useless	adjective	Feelings	negative

Among other frequently cited topic domains are *feelings* (22.7%) and *easy/hard* (20.1%). Most of the comments associated with these topics discuss the overall impression of the thermostat, such as STR #401 "*Love this product – works exceptionally well and looks terrific …*". The majority of reviews in these two topic domains were positive (79.6% and 71.5% respectively). This is not surprising given that the overall product ratings were between 4 and 5 out of 5 stars. Previous studies have suggested that the touch-screen interface of smart thermostats makes them more readable and easier to use [38]. This study's results in the *easy/hard* category confirm this. Other studies have indicated that some users find smart devices hard to understand and operate [39], however, this was not found in this dataset. This may be because now, more than 5 years after [39], smartphones, tablets and other similar electronic devices are more widespread, thus the average person is likely now more comfortable with such interfaces [40]. This may increase people's comfort level with thermostat interfaces compared to when they were first introduced to the market. The thermostat user interface design has also likely improved in contrast with the thermostat designs referred to in Ref. [39].

Connectivity related bi-terms make up more than 10% of all collected bi-terms in the smart thermostat database. Similar to the *control/occupancy* features, the majority of the associated bi-terms in this topic category are tagged either as positive (35.9%) or neutral (56.5%). This suggests that some users find the connectivity related features interesting and tend to comment on them positively.

For instance, the reviewers often assessed the device's connectivity with regards to its remote access function. As an example, STR #17848 stated: "*The feature that made the [device name] my choice was the WiFi connectivity that allows remote control of the home and away functions. I installed the [device name] in my cabin that is a 2 h drive from my home and has access only by boat. I can go into the app and turn the heat mode to home when I leave and have a warm cabin to enter when I arrive.*" However, a minority of reviewers (7.6%) expressed concerns, particularly with sudden or reoccurring loss of WiFi connection or incompatibility of the provided software with their existing digital infrastructure.

Approximately 10% of the comments focused on *installation.* Once again, the majority of the associated bi-terms are either positive (50.7%) or neutral (42.2%). A qualitative review shows that most of the reviewers discussed the devices' installation procedure. Some mentioned that despite their lack of previous knowledge and/or confidence in their ability to complete the installation, the installation was straightforward, and expressed that they were able to install without major difficulties. For example, STR #9000 stated: "*Superb installation and setup instructions. Obviously, a lot of thought went into ensuring that this product would be easy to install. Did it myself - a breeze …*" The customer/technical support from the provider during the installation is also frequently mentioned; for example, STR #641 stated that: "*Easy to install even for a novice – literally someone here who had never even seen the inside of a thermostat before. I followed the instructions (having also viewed their YouTube videos). I called tech support to ensure I had set things up properly. They asked me to email pics of my connections and instantly confirmed all*

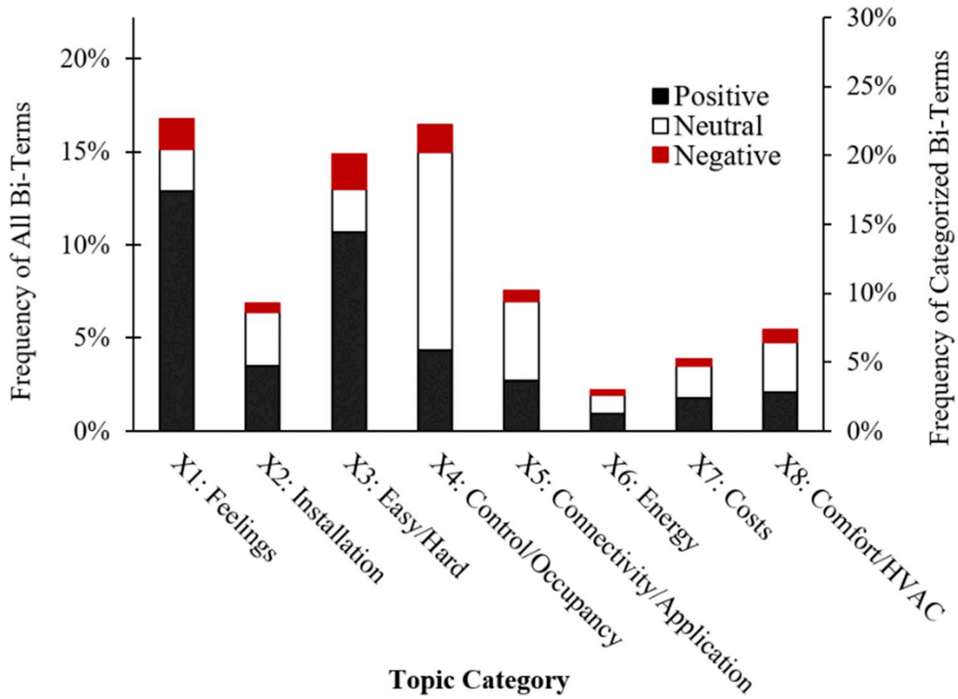

Fig. 3. Bi-term frequency per topic domain for all smart thermostat reviews.

was well. They even took me through their advanced system test to confirm the unit was functioning properly …" Based on these comments, it appears that the combination of comprehensible instructions and reliable technical support team are well received. Ease of installation and commissioning, therefore, may also be a metric of importance to potential customers.

The topic of *comfort* (7.4% of categorized bi-terms) was also discussed in some reviews. Although this topic was not as frequently discussed compared to other categories, it was brought up more than twice as much as the *energy* topic. This proportional frequency of appearance confirms the findings of Zipperer et al. (2013) who suggested users are more interested in their comfort and understand energy savings not in terms of using less energy, but as using the same amount of energy to gain more comfort [41]. For instance, STR #1794 and #426 respectively stated: "*Comfort was our goal. Energy savings was an extra bonus*." and "… *I will be interested to see how much it saves on energy, but if for nothing but the convenience it is well worth it*." This phenomenon of more energy use when more efficient/less energy consuming systems are used is often referred to as the "rebound effect", which has the potential to negate energy savings from more efficient systems, devices and/or configurations.

Energy is the least discussed topic category among the 8 topic domains, with less than 3% of all the collected bi-terms. The limited discussion of energy relative to other topics is somewhat surprising given that earlier studies have identified energy savings as being among the most important factors considered by smart thermostat purchasers [16,25,42]. This may point to a shift in factors motivating the purchasers and/or ultimately the users of smart thermostats. This difference may also be due in part to the inclusion of a broader range of smart thermostat users in this data who may value other thermostat features, as compared to earlier studies that utilized a population more strongly dominated by more energy-conscious early adopters [16]. It is also possible that there is an unspoken expectation to save energy, such as discussed in STR #2327 "… *while I don't get too much into the energy saving reports it produces, I do believe it maximizes the efficiency and minimizes my heating and cooling bill* …". Another point of consideration is that reviews were likely written before energy savings were verifiable. A few reviewers, however, did updated their review, such as STR #8860: "… *Update: Got my first utility bill since installing the thermostat. Am seeing savings already. At current rates, the thermostat will pay for itself inside a year* …". In summary, however, most users did not update their review after a longer period of use and evaluated the device based on its other non-energy related benefits and features. This may suggest that energy cannot be identified as an immediate concern and/or factor of consideration of the user.

Cost was also among the least-discussed topics. Previous studies have suggested that the higher cost and lower perceived "value" of smart thermostats were barriers to widespread adoption [41]. However, relative to other categories, this data suggests there is relatively little discussion on the monetary aspects of these devices (5.2% of bi-terms), and even when this is discussed, the sentiment is typically either positive or neutral (90.1%). This may be due in part to only having comments from people that had already decided to purchase a smart thermostat. In fact, in several cases reviewers indicated that the price was justified, such as STR #18889: "*You will not regret it. The savings alone are well worth the cost*." Many of the money related comments also include comparisons between different smart thermostats, such as STR #14101 and STR #6419, which state "… *Well worth the money spent when compared to others on the market in my opinion* …", and "… *Considering it is at least $50 cheaper than [another smart thermostat], it is well worth the money* …" This suggests that market trends and barriers to adoption based on this more recent datasets may be different than what previous literature had suggested.

A breakdown of the results by product (Fig. 4) rather than in aggregated form, provides information on the relative share of each topic category and sentiment. The data for each individual thermostat is generally similar to the overall dataset. This confirms that even though the number of collected reviews per device was inevitably somewhat different for each product based on the number of available reviews online, the results do not appear to be skewed by a single product's reviews. Further, it seems that users' perceptions towards smart thermostats are not necessarily impacted by the differences in features from one device to another. The reviewers generally discuss every topic with the same approximate frequency regardless of these differences. Moreover, the aggregated intensity of their opinion (in terms of being positive, negative or neutral) in each topic category also tends to remain

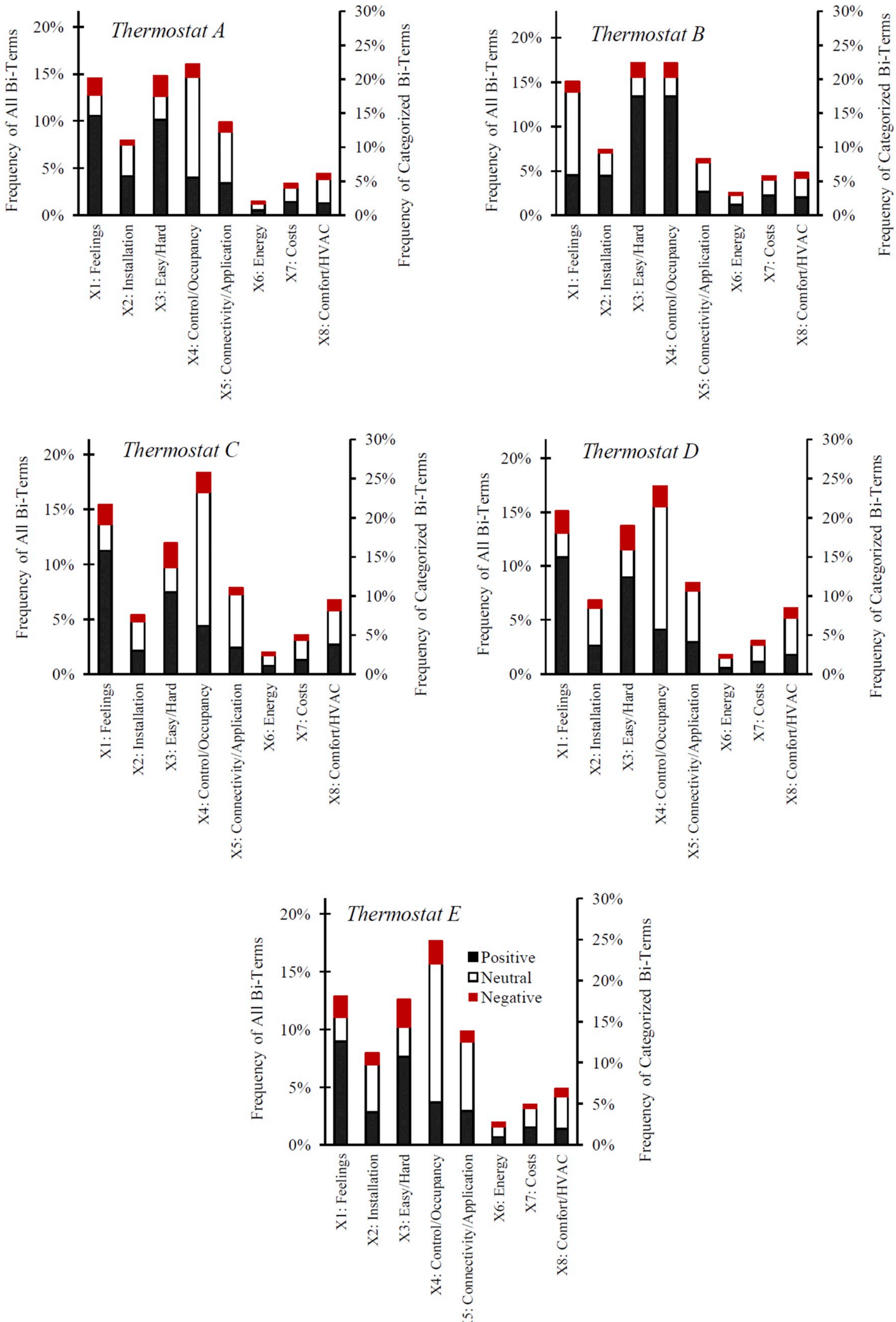

Fig. 4. Bi-term frequency by topic domain for the selected smart thermostats by product, for Thermostats A - E.

nearly the same across the different studied products.

3.2. Non-smart thermostat reviews

The data for the non-smart thermostats is noticeably different from that of the smart thermostats. Using the same topic categories developed for the smart thermostats, only approximately 28% of the collected bi-terms for non-smart thermostat database fit into the eight pre-defined topic categories. This low percentage supports the different thematic focuses in the two databases. Fig. 5 shows the appearance frequency of the collected bi-terms in the non-smart thermostat reviews database per topic domain.

Comparing the non-smart thermostat data (Fig. 5) to that of the smart thermostats (Fig. 3), the categories of bi-terms that vary the most are *control/occupancy* and *connectivity/application*. This is not surprising given that these two features generally represent some of the most commonly discussed advantages of smart thermostats over the non-smart ones. This is especially true in the case of the *connectivity/application* related features which are usually not included in the non-smart thermostats. In the case of *control/occupancy* related features, while the non-smart thermostats have some functionalities that enable programming of setpoint schedules, these capabilities are comparatively limited. Another difference between the two datasets is with regards to the *comfort/HVAC* topic domain. Although it is discussed relatively as often as with the non-smart thermostat reviews (7.5% compared to 7.4%), but in the case of non-smart thermostats this topic category mostly consists of neutral bi-terms (86.9%), which was not the case in the smart thermostat reviews (which consisted of 37.9% positive and 49.1% neutral bi-terms).

The *energy*-related bi-terms are among the least discussed topics (1.6%), similar to the smart thermostats. The majority of the bi-terms in this topic domain were neutral in sentiment (76.9%), which differs from the smart thermostat reviews (41.8% positive, 46.1% negative, 12.1% neutral). As mentioned previously, users appear to generally expect smart thermostats to save them energy, but this seems to be less the case for non-smart thermostats. For instance, PTR #351 and PTR #1662 respectively stated: "… *perhaps it will save a little in my energy cost.*", and "… *Not sure we are saving money*." From these reviews, it seems as if users are not as confident in the energy saving capabilities of non-smart thermostats. The overall high ratings of the non-smart thermostats in this study, however, further demonstrate that that energy-saving features and capabilities of a device may not be a high priority to the costumers.

Fig. 6 represents the same bi-term frequency results for each of the selected non-smart thermostats. As it was the case in the smart thermostats analysis, these generally follow the same pattern in terms of frequency as the overall non-smart thermostats. There do not appear to be any significant outlier reviews of any single product.

4. Discussion

Based on the discussed results, there are several themes that emerge, as follows.

4.1. Users generally do not discuss energy efficiency

This data shows that in contrast to earlier studies (e.g. Refs. [16,25,42]), smart thermostat users are not necessarily focused solely on energy-saving benefits. Instead, based on the frequency of bi-terms in other categories of topics discussed, users seem to be more interested in other features. This has several potential implications: First, increasing the penetration and use of smart thermostats may require a shift in focus around the non-energy benefits of these devices. This shift in focus is especially critical since some previous research has found that the marketing of energy efficient technologies with a sole focus on their environmental benefits, for some portions of the population, can have a negative effect on the resulting adoption rates due to the political views that some people associate with supporting environmental issues [23,43]. Therefore, highlighting other non-energy related features and capabilities more could be beneficial. Second, home energy management devices (HEMs) in general and smart thermostats in particular function in a different way than other energy efficiency upgrades. While conventional energy efficiency upgrades (e.g. insulation) do not necessarily require user engagement after installation, HEMs require feedback and behavioral adjustment with the user. This point, also highlighted by previous research (e.g. Ref. [36]), suggests that since users' lack of

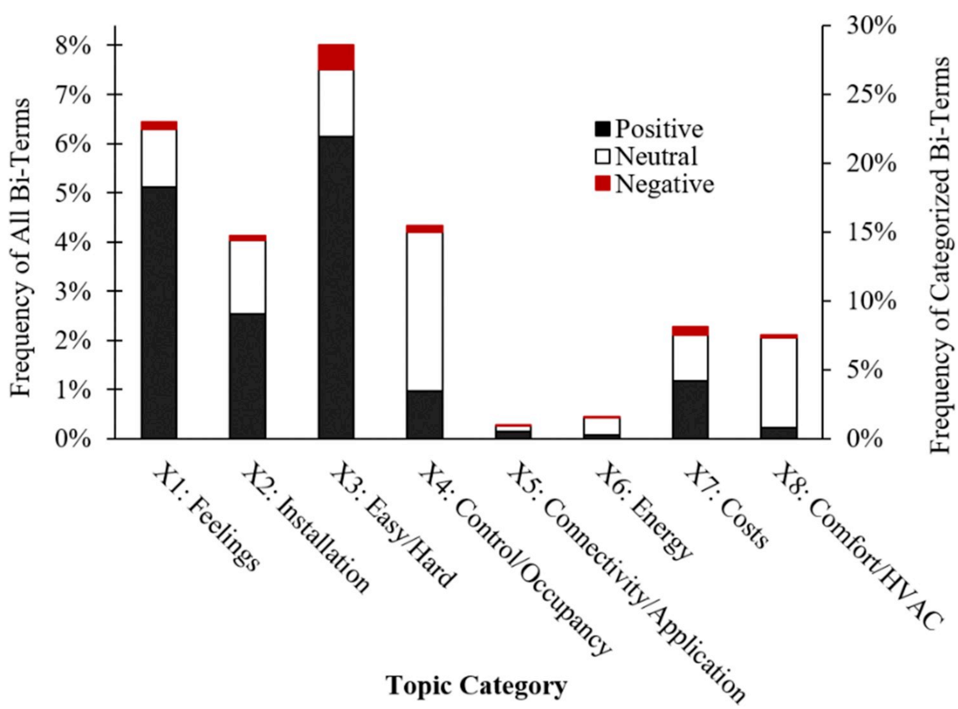

Fig. 5. Bi-term frequency per topic domain for all non-smart thermostat reviews.

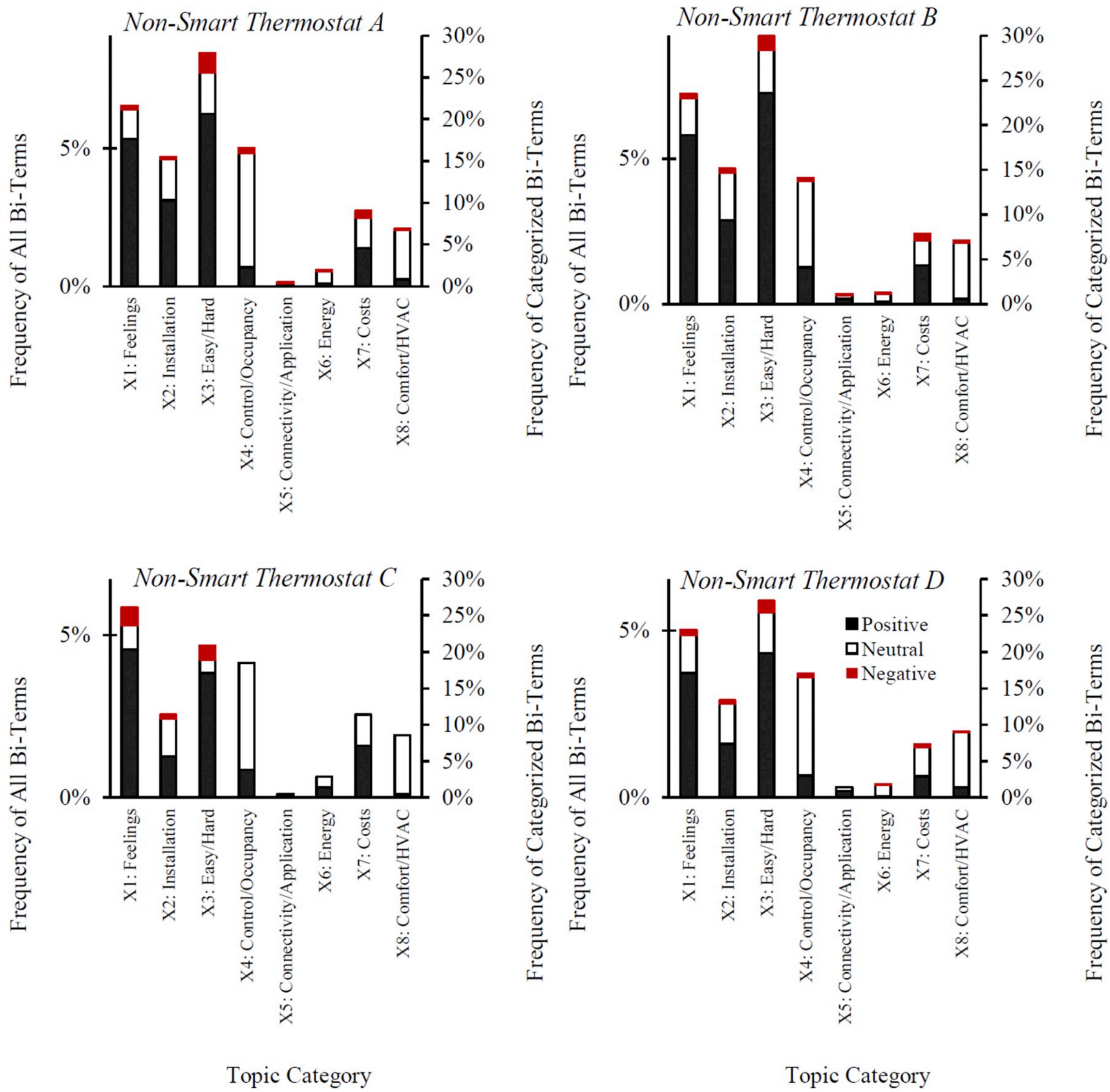

Fig. 6. Bi-term frequency by topic domain for non-smart thermostats A through D.

interest in energy savings can lead to low engagement level with the related features of their devices, the devices may not realize their energy-saving potential without more automated interventions. This is particularly important, since previous studies suggest that participants' engagement with the device generally decreases over time [16].

Regardless, a common barrier to use of energy efficient technology, including thermostats, is getting households to purchase and install such devices in the first place [44]. As such, if a focus on non-energy features increases further use of smart thermostats, the flexibility of such devices to engage households in energy saving controls or to automatically adjust HVAC use to save energy is beneficial, even if this was not the reason the household purchased the device.

4.2. Early adopters and the current user population may differ in interest and motivations

Most of the studies conducted on smart thermostats are focused on studying earlier adopters of the technology and their perception towards these devices [16,25]. However, the findings of this study differ when compared to the findings of such studies, specifically in that usability and upfront investment costs are no longer expressed as concerns. This suggests that values and interests of the user base may be changing, and would probably benefit from periodic re-evaluation. Shifts in users' perceptions, attitudes and behaviors towards any type of technology are inevitable, however, this can also be beneficial. For instance, high cost and low perceived value were previously identified as barriers to the adoption of smart thermostats [41,45]. The results of this work appears to suggest that this may not necessarily be the case currently.

4.3. Balancing control and automation

This study's results agree with the findings of previous research that suggest users' preferences with regards to control and/or automation are diverse. Therefore, it may be beneficial for the devices' control mechanisms to reflect this diversity in users' desires and provide a range of options for users to choose from Ref. [36]. These should include options from full automation to manual user-defined controls. As a result, a user would have the option to choose the most suitable setting for his/her preferences. However, given that most users will likely use the thermostat in the "default" mode, it may be beneficial to have the default mode be that which continues to support energy savings through some level of automation. In addition, prior research has also demonstrated that users' engagement with such technologies typically decreases over time [16], thus it may be beneficial to periodically attempt to engage the user regardless of their choice of settings. Finally, given that households are not likely to change their thermostat hardware often, it might be beneficial if, similar to smart phones and other electronic devices, smart thermostat software is able to be updated regularly. Accordingly, when improved features and control mechanisms are developed, these can be added to an installed device to facilitate more energy savings and enhance the user experience.

4.4. Technical support is a valued feature of the lived experience of smart thermostats

Regarding barriers to adoption and efficient use of smart home technologies, technical support is rarely mentioned. However, the users from this dataset commonly noted that technical support from the manufacturer was important to their experience with the device, particularly during troubleshooting. Some even went as far as referring to technical support as a "miracle-maker" and stated that if it was not for the availability of this service, they would not have been able to install and setup the device. This is particularly of importance, since prior studies have associated independent installation with improved (and increased) interactions with the device over the long-term [45]. As such, explaining to potential users that availability of technical support may help encourage both adoption and proper use.

4.5. Users seem to value comfort over energy efficiency

According to the results of this analysis, comfort appears to be of relatively larger concern to the user than energy efficiency. In this data, comfort was mentioned more than twice as much as energy. Previous research has found that users often expect that the device to automatically saves energy and do not necessarily understand how their choices can easily offset any potential savings [39]. Therefore, increasing user awareness and appreciation for a balance between comfort and energy savings may be beneficial. As demonstrated by prior work, real-time feedback is more beneficial than periodical history reports since they are more likely to trigger behavioral change [46].

4.6. Usability is not an expressed concern

It appears that a combination of higher technological literacy and better interface design may have contributed to reduced concerns and negative commentary on the usability and installation issues that early users had expressed in the previous studies (e.g. Ref. [39]). In this work, reviewers generally indicated they were comfortable and confident in their abilities to install (only 7.1% negative), setup and monitor (only 8.9% negative) their device on their own and had an overall positive experience with the device.

5. Conclusions

In this study, online reviews were used to evaluate users' perceptions and attitudes towards smart thermostats. For this purpose, 26,372 reviews posted on an online retailer's website (Amazon.com) were collected for five commercially available popular smart thermostats. These reviews were then analyzed using a confirmatory aspect-based opinion mining technique. Each review's text was summarized using bi-terms that each included an aspect word and an opinion word. These bi-terms were subsequently categorized into eight topic domains and assigned a sentiment score of "positive, "negative", or "neutral". Analysis of the results demonstrated that "energy" was surprisingly the least discussed topic among review categories, which contrasts with previous research which had suggested that energy saving was the main motivational factor involved in the purchase of smart thermostats [16,25]. This may be because early adopters are not entirely representative of the current user population, or that perhaps the different nature of the means of data collection (i.e surveys and laboratory testing versus online reviews) may have provided a different dataset of user feedback. In this regard, online reviews appear to provide a unique opportunity for researchers to investigate users' perceptions toward emerging technology with no influence on their responses.

The results of this work agree with previous research that suggested users are more concerned about comfort than energy efficiency. Smart home energy management technologies are fundamentally different than other conventional energy efficiency upgrades (e.g. insulation) in that they require regular interactions from the user. Therefore, current and future users must be educated on this difference and learn about the consequences of their everyday choices.

With regards to ease of use and installation, previously identified barriers were not mentioned by the large majority of reviewers, which indicated that they were confident in their ability to install, setup, and program their devices on their own. Upgraded interface designs and higher digital skills among a broader user base are both likely to have contributed to this observed enhancement in user experience. In the specific case of initial installation, real-time technical support was mentioned numerous times and users noted that this support highly improved their experience.

Lastly, this analysis revealed conflicting ideas about the issues of control and/or autonomy. While some users were happy to surrender control, others were reluctant to allow for automation, which suggests a need for diversity in control options for users to choose from. However, this should be balanced with the need to continue to strive towards utilizing the thermostat to support the most energy efficient control decisions.

Overall, by evaluating users' perceptions and attitudes towards smart thermostats, the findings of this study may be beneficial to manufacturers of smart thermostats who may be interested in better understanding their customer and user base and their associated sentiments and opinions about thermostat products. This may also be of interest to those that design and execute energy efficiency rebate programs, including those with smart thermostats, to understand what features are of interest and associated with positive sentiment to potential customers interested in participating in such programs. Limitations of this work are first related to the unavailability of a field-specific sentiment dictionary. The creation and use of a field-specific sentiment dictionary, as compared to the generic sentiment dictionary used in this work, could help capture thermostat-specific and -related words and their associated sentiments that otherwise may not be included and/or designated in a standard sentiment dictionary. A second limitation is the exclusion of input data from potential buyers from the dataset used for analysis. Given that those that do not purchase thermostats do not write online reviews which could be used in this dataset, it is unlikely that a parallel dataset of potential users' reviews of thermostats exists. Future work would benefit from a comparative side-by-side analysis of both current and potential users. Additionally, the analysis presented in this manuscript is following the assumptions made in other related work [47,48] that correlate the frequency of user's reported sentiments in online reviews directly with the values they put on product features. Future work may investigate this further to provide insight into its applicability for similar studies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The fourth author, Jongho Im's, research is supported by National Research Foundation of (NRF) Korea, NRF-2018R1D1A1B07045220.

References

[1] L. Doman, EIA projects 28% increase in world energy use by 2040, U.S. Energy Inf. Adm., 2017.
[2] U.S. Energy Information Administration, Annual Energy Outlook 2018 with Projections to 2050, 2018. https://www.eia.gov/outlooks/aeo/pdf/AEO2018.pdf.
[3] U.S. Energy Information Administration, Annual Energy Outlook 2019 with Projections to 2050, 2019. https://www.eia.gov/outlooks/aeo/pdf/aeo2019.pdf.
[4] C. Diakaki, E. Grigoroudis, D. Kolokotsa, Towards a multi-objective optimization approach for improving energy efficiency in buildings, Energy Build. 40 (2008) 1747–1754.
[5] H. Do, K.S. Cetin, Data-driven evaluation of residential HVAC system efficiency using energy and environmental data, Energies 12 (2019) 188.
[6] B.L.R. Stojkoska, K. V Trivodaliev, A review of Internet of Things for smart home: challenges and solutions, J. Clean. Prod. 140 (2017) 1454–1464.
[7] C. Ryan, Computer and Internet Use in the United States: 2016, 2018. https://www.census.gov/content/dam/Census/library/publications/2018/acs/ACS-39.pdf.
[8] Statista, Smart Home (United States). https://www.statista.com/outlook/279/109/smart-home/united-states, 2019.
[9] U.S. Energy Information Administration, Energy Use in Homes, 2019.
[10] M. Avci, M. Erkoc, A. Rahmani, S. Asfour, Model predictive HVAC load control in buildings using real-time electricity pricing, Energy Build. 60 (2013) 199–209.
[11] P. Wattles, Demand response and ERCOT grid reliability, in: Energy Effic. Forum e Arlington, TX, 2012.
[12] Houzz Research, U.S. Houzz Smart Home Trends Study, 2016, 2017, https://www.houzz.com/magazine/2016-us-houzz-smart-home-trends-study-stsrsivw-vs~92045400.
[13] A. Bardhan, D. Jaffee, C. Kroll, N. Wallace, Energy efficiency retrofits for US housing: removing the bottlenecks, Reg. Sci. Urban Econ. 47 (2014) 45–60.
[14] U.S. Department, Of Energy, Overview of Existing and Future Residential Use Cases for Connected Thermostats, 2016.
[15] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, K. Whitehouse, The smart thermostat: using occupancy sensors to save energy in homes. Proc. 8th ACM Conf. Embed. Networked Sens. Syst., ACM, 2010, pp. 211–224.
[16] R. Yang, M.W. Newman, Learning from a learning thermostat: lessons for intelligent systems for the home, in: Proc. 2013 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput., ACM, 2013, pp. 93–102.
[17] American Council for an Energy-Efficient Economy, Energy Impacts of Smart Home Technologies, 2018. https://aceee.org/sites/default/files/publications/researchreports/a1801.pdf.
[18] U.S. Energy Information Administration, RECS Survey Data, 2015, 2015, https://www.eia.gov/consumption/residential/data/2015/.
[19] A. Darko, A.P.C. Chan, Review of barriers to green building adoption, Sustain. Dev. 25 (2017) 167–179.
[20] A.P.C. Chan, A. Darko, E.E. Ameyaw, D.-G. Owusu-Manu, Barriers affecting the adoption of green building technologies, J. Manag. Eng. 33 (2016) 4016057.
[21] M. Ringel, R. Laidi, D. Djenouri, Multiple benefits through smart home energy management solutions—a simulation-based case study of a single-family-house in Algeria and Germany, Energies 12 (2019) 1537.
[22] E. Mills, A. Rosenfeld, Consumer non-energy benefits as a motivation for making energy-efficiency improvements, Energy 21 (1996) 707–720.
[23] J. Im, Y. Seo, K.S. Cetin, J. Singh, Energy efficiency in US residential rental housing: adoption rates and impact on rent, Appl. Energy 205 (2017) 1021–1033.
[24] E.R. Frederiks, K. Stenner, E. V Hobman, Household energy use: applying behavioural economics to understand consumer decision-making and behaviour, Renew. Sustain. Energy Rev. 41 (2015) 1385–1394.
[25] R. Yang, M.W. Newman, Living with an intelligent thermostat: advanced control for heating and cooling systems, in: Proc. 2012 ACM Conf. Ubiquitous Comput., ACM, 2012, pp. 1102–1107.
[26] E. Marrese-Taylor, J.D. Velásquez, F. Bravo-Marquez, A novel deterministic approach for aspect-based opinion mining in tourism products reviews, Expert Syst. Appl. 41 (2014) 7764–7775.
[27] K. Schouten, F. Frasincar, Survey on aspect-level sentiment analysis, IEEE Trans. Knowl. Data Eng. 28 (2015) 813–830.
[28] P.P. Tribhuvan, S.G. Bhirud, A.P. Tribhuvan, A peer review of feature based opinion mining and summarization, IJCSIT), Int. J. Comput. Sci. Inf. Technol. 5 (2014) 247–250.
[29] I. Jongho, T. Song, Y. Lee, K. Jewoo, Confirmatory aspect-based opinion mining processes, arXiv, 2019 arXiv:1907.12850.
[30] H. Wickham, M.H. Wickham, Package 'rvest', 2019.
[31] R.C. Team, R, A Language and Environment for Statistical Computing, 2013.
[32] I. Feinerer, Introduction to the tm package text mining in R, access, En Ligne (2013). http://Cran.r-Project.Org/Web/Packages/Tm/Vignettes/Tm.Pdf.
[33] M. Hu, B. Liu, Mining and summarizing customer reviews. Proc. Tenth ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., ACM, 2004, pp. 168–177.
[34] J. Silge, D. Robinson, Text Mining with R: A Tidy Approach, O'Reilly Media, Inc., ", 2017.
[35] G.A. Miller, WordNet: a lexical database for English, Commun. ACM 38 (1995) 39–41.
[36] T. Peffer, M. Pritoni, A. Meier, C. Aragon, D. Perry, How people use thermostats in homes: a review, Build. Environ. 46 (2011) 2529–2541.
[37] T.E. Peffer, California DREAMing: the Design of Residential Demand Responsive Technology with People in Mind, 2009.
[38] A. Meier, C. Aragon, B. Hurwitz, D. Mujumdar, T. Peffer, D. Perry, M. Pritoni, How People Actually Use Thermostats, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2010.
[39] L.C.R. Oliveira, A. May, V. Mitchell, M. Coleman, T. Kane, S.K. Firth, Pre-installation Challenges: Classifying Barriers to the Introduction of Smart Home Technology, 2015.
[40] J. Horiggan, Digital Readiness Gaps, 2016. https://www.pewinternet.org/wp-content/uploads/sites/9/2016/09/PI_2016.09.20_Digital-Readiness-Gaps_FINAL.pdf.
[41] A. Zipperer, P.A. Aloise-Young, S. Suryanarayanan, R. Roche, L. Earle, D. Christensen, P. Bauleo, D. Zimmerle, Electric energy management in the smart home: perspectives on enabling technologies and consumer behavior, Proc. IEEE 101 (2013) 2397–2408.
[42] R. Yang, M.W. Newman, J. Forlizzi, Making sustainability sustainable: challenges in the design of eco-interaction technologies, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM, 2014, pp. 823–832.
[43] D.M. Gromet, H. Kunreuther, R.P. Larrick, Political ideology affects energy-efficiency attitudes and choices, Proc. Natl. Acad. Sci. 110 (2013) 9314–9319.
[44] J.G. Koomey, C.A. Webber, C.S. Atkinson, A. Nicholls, Addressing energy-related challenges for the US buildings sector: results from the clean energy futures study, Energy Policy 29 (2001) 1209–1221.
[45] A.J. Brush, B. Lee, R. Mahajan, S. Agarwal, S. Saroiu, C. Dixon, Home automation in the wild: challenges and opportunities, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM, 2011, pp. 2115–2124.
[46] S. Houde, A. Todd, A. Sudarshan, J.A. Flora, K.C. Armel, Real-time feedback and electricity consumption: a field experiment assessing the potential for savings and persistence, Energy J. (2013) 87–102.
[47] B. Liu, M. Hu, J. Cheng, Opinion observer: analyzing and comparing opinions on the web, in: Proc. 14th Int. Conf. World Wide Web, ACM, 2005, pp. 342–351.
[48] B. Liu, Sentiment analysis and subjectivity, Handb. Nat. Lang. Process. 2 (2010) 627–666.