A Data-Driven Approach to Developing IoT Privacy-Setting Interfaces

Paritosh Bahirat*, Yangyang He*, Abhilash Menon*, Bart Knijnenburg
Clemson University School of Computing
Clemson, USA
{pbahira,yyhe,abhilas,bartk}@clemson.edu

ABSTRACT

User testing is often used to inform the development of user interfaces (UIs). But what if an interface needs to be developed for a system that does not yet exist? In that case, existing datasets can provide valuable input for UI development. We apply a data-driven approach to the development of a privacy-setting interface for Internet-of-Things (IoT) devices. Applying machine learning techniques to an existing dataset of users' sharing preferences in IoT scenarios, we develop a set of "smart" default profiles. Our resulting interface asks users to choose among these profiles, which capture their preferences with an accuracy of 82%—a 14% improvement over a naive default setting and a 12% improvement over a single smart default setting for all users.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous;

Author Keywords

Data-driven design; Internet of Things; Privacy settings; Machine learning

INTRODUCTION

Under the moniker of 'Internet of Things' (IoT), smart connected devices are revolutionizing our everyday life, just like smartphones did for cellphones. Smartphones, however, have shown to increase users' privacy concerns [4], and the same may be true for IoT. Like smartphones, IoT devices collect and store personal information to personalize the user experience, share it across other devices, and/or sell it to third parties. Consequently, preserving users' privacy is a big concern that limits the adoption of IoT devices [10].

*: These authors contributed equally.

IUI 2018, March 7–11, 2018, Tokyo, Japan.

http://dx.doi.org/10.1145/3172944.3172982

Privacy is an inherent trade-off in IoT, because IoT devices cannot provide their services without collecting data. Preserving users' privacy therefore means giving them control over this trade-off, by allowing them to decide what information can be collected about them. Outside the home environment, people have little control over the data IoT devices collect. Researchers at Intel are working on a framework that allows people to be notified about surrounding IoT devices collecting personal information, and to control these collection practices [5].

Smartphones give users control over their privacy settings in the form of prompts that ask whether the user allows or denies a certain app access to a certain type of information. Such prompts are problematic for IoT, because IoT devices are supposed to operate in the background. Moreover, as the penetration of IoT devices in our environment continues to increase, prompts would become a constant noise which users will soon start to ignore, like software EULAs [8] or privacy policies [12].

A better solution would be to regulate privacy with global settings. But research has shown that users are highly concerned about their privacy, but find it difficult to implement privacy settings [1, 9, 19]. Indeed, the vast number of encounters people have with a myriad of different IoT devices makes chosing adequate privacy settings a very challenging task that is likely to result in information and choice overload [28].

Data-driven design

What design process allows us to develop a usable privacy-setting interface for IoT? The development of usable privacy interfaces commonly relies on user studies with existing systems. However, this method is not possible in our IoT control scenario, because the Intel control framework has yet to be implemented [5]. We therefore develop and employ a *data-driven design* methodology, leveraging an existing dataset collected by Lee and Kobsa [16], who asked users whether they would allow or deny IoT devices in their environment to collect information about them. We use this dataset in two phases.

In our first phase, we develop a "layered" settings interface, where users make a decision on a less granular level (e.g., whether a certain recipient is allowed to collect their personal information or not), and only move

to a more granular decision (e.g., what types of information this recipient is allowed to collect) when they desire more detailed control. This reduces the complexity of the decisions users have to make, without reducing the amount of control available to them. We use statistical analysis of the Lee and Kobsa dataset to decide which aspect should be presented at the highest layer of our IoT privacy-setting interface, and which aspects are relegated to subsequently lower layers.

In our second phase, we develop a "smart" default setting, which preempts the need for many users to manually change their settings [26]. However, since people differ extensively in their privacy preferences [20], it is not possible to achieve an optimal default that is the same for everyone. Instead, different people may require different settings. Outside the field of IoT, researchers have been able to establish distinct clusters or "profiles" based on user behavioral data [14, 20, 29]. We perform machine learning analysis on the Lee and Kobsa dataset to create a similar set of "smart profiles" for our IoT privacy-setting interface.

The remainder of this paper is structured as follows: We first summarize previous work on privacy in IoT scenarios, and describe the structure of the Lee and Kobsa [16] dataset. We then *inspect* users' behaviors using statistical analysis. Next, we *predict* users' behaviors using machine learning methods. We subsequently present a set of prototypes for an IoT privacy-setting interface. Finally, we conclude with a summary of our proposed procedure and the results of our analysis.

APPROACH AND RELATED WORK

Our goal is to develop intuitive interfaces for IoT privacy settings, using a data-driven approach. In this section we therefore discuss existing research on privacy-setting interfaces and on privacy prediction.

Privacy-Setting Interfaces

The most basic privacy-setting interface is the "access control matrix", which allows users to indicate who gets to see what [25]. This can be simplified by grouping recipients into categories, such as Google+'s *circles* [27]. Taking a step further, Raber et al. [22] proposed *Privacy Wedges*, which allow users to make privacy decisions using a combination of categorization (the wedges) and inter-personal distance (the position of a person on the wedge). Users can decide who gets to see their posts or personal information by "coloring" parts of each wedge.

These wedges have been tested on limited numbers of friends, and in the case of IoT they are likely insufficient, due to the complexity of the decision space. To wit, IoT privacy decisions involve a large selection of devices, each with various sensors that collect data for a range of different purposes. This makes it complicated to design an interface that covers every possible setting [28]. A wedge-based interface will arguably not be able to succinctly represent such complexity without a significant amount of information and choice overload.

We propose a data-driven approach to solve this problem: statistical analysis informs the construction of a layered settings interface, while machine learning-based privacy prediction helps us find smart privacy profiles.

Privacy Prediction

Several researchers have proposed privacy prediction as a solution to the privacy settings complexity problem. Sadeh et al. used a k-nearest neighbor algorithm and a random forest algorithm to predict users' privacy preferences in a location-sharing system [24], based on the type of recipient and the time and location of the request. They demonstrated that users had difficulties setting their privacy preferences, and that the applied machine learning techniques can help users to choose more accurate disclosure preferences. Similarly, Pallapa et al. [21] present a system which can determine the required privacy level in new situations based on the history of interaction between users. Their system can efficiently deal with the rise of privacy concerns and help users in a pervasive system full of dynamic interactions.

Dong et al. [6] use a binary classification algorithms to give users personalized advice regarding their privacy decision-making practices on online social networks. They found that J48 decision trees provided the best results. Li and et al. [17] similarly use J48 to demonstrate that taking the user's cultural background into account when making privacy predictions improves the prediction accuracy. Our data stems from a culturally homogeneous population (U.S. Mechanical Turk workers), so cultural variables are outside the scope of our study. We do however follow these previous works in using J48 decision trees in our prediction approach.

We further extend our approach using *clustering* to find several smart default policies ("profiles"). This is in line with Fang et al. [7], who present an active learning algorithm that comes up with privacy profiles for users in real time. Since our approach is based on an existing dataset, our algorithm does not classify users in real time, but instead creates a static set of profiles 'offline', from which users can subsequently choose. This avoids cold start problems, and does not rely on the availability of continuous real-time behaviors. This is beneficial for IoT settings, because users often specify their settings in these systems in a "single shot", leaving the settings interface alone afterwards.

Ravichandran et al. [23] employ an approach similar to ours, using k-means clustering on users' contextualized location sharing decisions to come up with several default policies. They showed that a small number of default policies could accurately reflect a large part of the location sharing preferences. We extend their approach to find the best profiles based on various novel clustering approaches, and take the additional step of designing user interfaces that incorporate the best solutions.

We apply our procedure to a dataset by Lee and Kobsa [16], who presented users with a total of 2800

IoT usage scenarios that were systemstically manipulated along five dimensions. Lee and Kobsa observed that the scenarios can be grouped into four clusters in terms of privacy risks. Their clusters differ substantially along several dimensions, most notably regarding the inquirer ('who') and data type ('what'). The dominance of the 'who' parameter is also reflected in a study in a ubiquitous computing environment by Lederer et al. [15]. Extending upon Lee and Kobsa, our clustering procedure is performed at the *user* level rather than the *scenario* level. This allows us to create privacy profiles.

DATASET

This study is based on a dataset collected by Lee and Kobsa [16]. A total of 2800 scenarios were presented to 200 participants (100 male, 99 female, 1 undisclosed) through Amazon Mechanical Turk. Four participants were aged between 18 and 20, 75 aged 20–30, 68 aged 30–40, 31 aged 40–50, 20 aged 50–60, and 2 aged > 60.

Each participant was presented with 14 scenarios describing a situation where an IoT device would collect information about the participant. Each scenario was a combination of five contextual parameters (Table 1), manipulated at several levels using a mixed fractional factorial design that allowed us to test main effects and two-way interactions between all parameters.

For every scenario, participants were asked a total of 9 questions. Our study focuses on the **allow/reject** question: "If you had a choice to allow/reject this, what would you choose?", with options "I would allow it" and "I would reject it". We also used participants' answers to three attitudinal questions regarding the scenario:

- **Risk:** How risky or safe is this situation? (7pt scale from "very risky" to "very safe")
- **Comfort:** How comfortable or uncomfortable do you feel about this situation? (7pt scale)
- **Appropriateness:** How appropriate do you consider this situation? (7pt scale)

INSPECTING USERS' BEHAVIORS

In this section we analyze how users' behavioral intentions to allow or reject the information collection described in the scenario are influenced by the scenario parameters. In line with classic attitude-behavior models [2], we also investigate whether users' attitudes regarding the scenario—their judgment of risk, comfort, and appropriateness—mediate these effects. This mediation analysis [3] involves the following test:

- **Test 1:** The effect of the scenario parameters (who, what, where, reason, persistence) on participants' attitudes (risk, comfort, appropriateness).
- **Test 2:** The effect of participants' attitudes on their behavioral intentions (the allow/reject decision).
- **Test 3:** The effect of the parameters on behavioral intentions, controlling for attitudes.

Table 1: Parameters used in the experiment. Example scenarios:
"A device of a friend records your video to detect your presence. This happens continuously, while you are at someone else's place, for your safety."
"A government device reads your phone ID to detect your identity. This happens once, while you are in a public place (e.g. on the street), for health-related purposes."

Parameter	**Levels**
Who *The entity collecting the data*	1. Unknown 2. Colleague 3. Friend 4. Own device 5. Business 6. Employer 7. Government
What *The type of data collected and (optionally) the knowledge extracted from this data*	1. PhoneID 2. PhoneID>identity 3. Location 4. Location>presence 5. Voice 6. Voice>gender 7. Voice> age 8. Voice>identity 9. Voice>presence 10. Voice>mood 11. Photo 12. Photo>gender 13. Photo>age 14. Photo>identity 15. Photo>presence 16. Photo>mood 17. Video 18. Video>gender 19. Video>age 20. Video>presence 21. Video>mood 22. Video>looking at 23. Gaze 24. Gaze>looking at
Where *The location of the data collection*	1. Your place 2. Someone else's place 3. Semi-public place (e.g. restaurant) 4. Public space (e.g. street)
Reason *The reason for collecting this data*	1. Safety 2. Commercial 3. Social-related 4. Convenience 5. Health-related 6. None
Persistence *Whether data is collected once or continuously*	1. Once 2. Continuously

Table 2: Effect of scenario on attitudes. Each model builds upon and is tested against the previous.

Model	χ^2	*df*	*p*-value
$risk \sim (1\|sid)$			
+who	315.37	6	$< .0001$
+what	67.74	23	$< .0001$
+reason	15.65	5	.0079
+persistence	9.95	1	.0016
+where	7.47	3	.0586
+who:what	166.47	138	.0050
Model	χ^2	*df*	*p*-value
$comfort \sim (1\|sid)$			
+who	334.06	6	$< .0001$
+what	83.24	23	$< .0001$
+reason	18.68	5	.0022
+persistence	14.73	1	.0001
+where	3.25	3	.3544
+who:what	195.07	138	.0001
Model	χ^2	*df*	*p*-value
$appropriateness \sim (1\|sid)$			
+who	315.77	6	$< .0001$
+what	72.87	23	$< .0001$
+reason	23.27	5	.0003
+persistence	8.97	1	.0027
+where	5.46	3	.1411
+who:what	214.61	138	$< .0001$

If tests 1 and 2 are significant, and test 3 reveals a substantial reduction in conditional direct effect (compared to the marginal effect), then we can say that the effects of the scenario parameters on participants' behavioral intention are mediated by their attitudes. Moreover, if the conditional direct effect is (close to) zero, then the effects are fully (rather than partially) mediated.

Scenario Parameters and Attitude

ANOVA Test of Main Effects

To understand the effect of the scenario parameters on participants' attitudes, we created a separate *linear mixed effects regression* (*lmer*) model with a random intercept (to account for repeated measures on the same participant) for each dependent variable (risk, comfort, appropriateness), using the scenario parameters as independent variables. We employed a forward stepwise procedure, adding the strongest remaining parameter into the model at each step and comparing it against the previous model. Table 2 shows that all parameters except **where** have a significant effect on each of the attitudes.

Post-hoc Comparisons

We also conducted Tukey post hoc analyses to better understand how the various values of each parameter influenced the attitudes. **Where** was excluded from these analyses, as it did not have an overall significant effect. Some key findings of these post hoc analyses are:

Who: Participants perceive more *risk* when the recipient of the information is 'unknown' than for any other recipient (d range = [0.640, 1.450] and all ps $< .001$, except for 'government': $d = 0.286$, $p < .05$). 'Government' is the next most risky recipient (d range = [0.440, 1.190], all ps $< .001$). Participants consider their 'own device' the least risky (d range = [0.510, 1.450], all ps $< .001$). Similar patterns were found for *comfort* and *appropriateness*.

Reason: Participants were more *comfortable* disclosing information for the purpose of 'safety' than for any other reason except 'health' (d range = [0.230, 0.355], all ps $< .05$). They also believe that disclosing information for the purpose of 'health' or 'safety' is more *appropriate* than for 'social' or 'commercial' purposes (d range = [0.270, 0.310], all ps $< .05$).

Persistence: Participants were more *comfortable*, found it more *appropriate*, and less *risky* to disclose their information 'once' rather than 'continuously' ($d = 0.146$, $p < .01$).

What: This parameter has a large number of values, so we decided to selectively test planned contrasts instead of post-hoc tests. We first compared different mediums (voice, photo, video) regardless of what is being inferred:

- Participants were significantly more *comfortable* with 'voice' than 'video' ($d = 0.260$, $p = .005$), and found 'voice' less *risky* ($d = -0.239$, $p = .005$) and more *appropriate* ($d = 0.217$, $p = .015$) than 'video'.
- Participants were significantly more *comfortable* with 'voice' than 'photo' ($d = 0.201$, $p = .007$) and found 'voice' more *appropriate* than 'photo' ($d = 0.157$, $p = .028$). There was no significant difference in terms of *risk* ($p = .118$).
- No differences were found between 'photo' and 'video' in terms of *risk* ($p = .24$), *comfort* ($p = .35$) and *appropriateness* ($p = .26$).

We also compared different inferences (e.g. age, gender, mood, identity) across mediums. The following planned contrasts were significant (all others were not):

- Participants were significantly more *comfortable* ($d = 0.363$, $p = .028$) and found it more *appropriate* ($d = 0.371$, $p = .018$) to reveal their 'age' rather than their 'identity'.
- Participants were significantly more *comfortable* ($d = 0.363$, $p = .008$) and found it more *appropriate* ($d = 0.308$, $p = .024$) to reveal their 'presence' rather than their 'identity'.

Interaction effects

We also checked for two-way interactions between the scenario parameters. The only significant interaction effect observed was between **who** and **what**. The last line of each section in Table 2 shows the results of adding this interaction to the model. Due to space concerns, we choose not to address the post-hoc analysis of the $7 * 24 = 168$ specific combinations of who and what.

Table 3: Effect of attitudes and scenario on allow/reject.

Model	OR	χ^2	*df*	*p*-value
$allow \sim (1\|sid)$				
+risk	0.25	1005.24	1	$< .0001$
+comfort	5.04	723.27	1	$< .0001$
+appropriateness	3.47	128.17	1	$< .0001$
+who		8.80	6	.1851
+what		26.07	23	.2976
+reason		19.33	5	.0017
+persistence		12.69	1	.0004

Table 4: Effect of scenario on allow/reject, *not* controlling for attitudes.

Model	χ^2	*df*	*p*-value
$allow \sim (1\|sid)$			
+who	221.36	6	$< .0001$
+what	78.55	23	$< .0001$
+reason	21.95	5	.0005
+persistence	20.64	1	$< .0001$

Attitude and Behavioral intention

To test the effects of participants' attitudes on their allow/reject decision, we ran a *generalized linear mixed effects regression* (*glmer*) with a random intercept and a logit link function to account for the binary dependent variable. We found significant effects of all the three attitudes on participants' allow/reject decision (see Table 3). Each 1-point increase in **risk** results in a 4.04-fold decrease in the odds that the scenario will be allowed ($p < .0001$). Each 1-point increase in **comfort** results in a 5.04-fold increase ($p < .0001$), and each 1-point increase in **appropriateness** results in a 3.47-fold increase ($p < .0001$).

Mediation Analysis

The bottom half of Table 3 shows the *conditional* effects of the significant parameters (who, what, reason, persistance) on participants' allow/reject decision, controlling for attitude. **Who** and **what** are no longer significant; these effects are thus fully mediated by attitude. The effects of **reason** and **persistance** are still significant, but smaller than the marginal effects (i.e., without controlling for attitude, see Table 4)—their χ^2s are reduced by 12% and 39%, respectively. This means that the mediation effect was substantial in all cases. The final mediation model is displayed in Figure 1.

Discussion of Statistical Results

Our statistical results show several patterns that can inform the development of an IoT privacy-setting interface. We find that **who** is the most important scenario parameter, and should thus end up at the top layer of our interface. People are generally concerned about IoT scenarios involving unknown and government devices, but less concerned about about data collected by their own devices. Mistrust of government data collection is in line with Li et al.'s finding regarding US audiences [17].

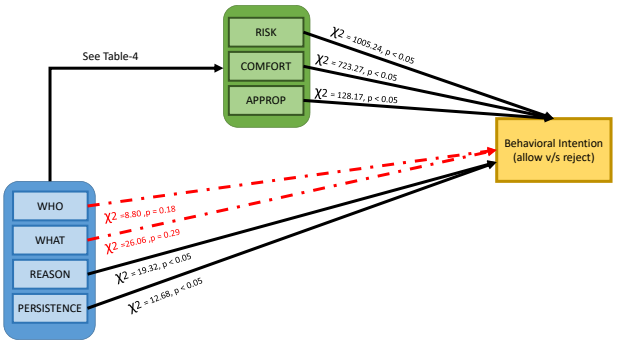

Figure 1: Mediation model of the effect of scenario parameters on participants' intention to allow/reject the scenario, mediated by attitudinal factors

What is the next most important scenario parameter, and its significant interaction with **who** suggests that some users may want to allow/reject the collection of different types of data by different types of recipients. Privacy concerns are higher for photo and video than for voice, arguably because photos and videos are more likely to reveal the identity of a person. Moreover, people are less concerned with revealing their age and presence, and most concerned with revealing their identity.

The **reason** for the data collection may be used as the next layer in the interface. Health and safety are generally seen as acceptable reasons. **Persistence** is less important, although one-time collection is more acceptable than continuous collection. **Where** the data is being collected does not influence intention at all. This could be an artifact of the dataset: location is arguably less prominent when reading a scenario than it is in real life.

Finally, participants' attitudes significantly (and in some cases fully) mediated the effect of scenario parameters on behavioral intentions. This means that these attitudes may be used as a valuable source for classifying people into distinct groups. Such attitudinal clustering could capture a significant amount of the variation in participants in terms of their preferred privacy settings, espcially with respect to the **who** and **what** dimensions.

PREDICTING USERS' BEHAVIORS

In this section we predict participants' allow/reject decision using machine learning methods. Our goal is to find suitable *default settings* for an IoT privacy-setting interface. Consequently, we do not attempt to find the best possible solution; instead we make a conscious tradeoff between parsimony and prediction accuracy. Accuracy is important to ensure that users' privacy preferences are accurately captured and/or need only few manual adjustments. Parsimony, on the other hand, prevents overfitting and promotes fairness: we noticed that more complex models tended to increase overall accuracy by predicting a few users' preferences more accurately, with no effect on other users. Parsimony also makes the associated default setting easier to understand for the user.

Table 5: Comparison of clustering approaches

Approach	clusters	Accuracy	# of profiles
Naive classification	1	28.33%	1 (all 'yes')
	1	71.67%	1 (all 'no')
Overall	1	73.10%	1
Attitude-based clustering	2	75.28%	2
	3	75.17%	3
	4	75.60%	3
	5	75.25%	3
Fit-based clustering	2	77.99%	2
	3	81.54%	3
Agglomerative clustering	200	78.13%	4
	200	78.27%	5

Our prediction target is the participants' decision to allow or reject the data collection described in each scenario, classifying a scenario as either 'yes' or 'no'. The scenario parameters serve as input attributes. These are nominal variables, making decision tree algorithms such as ID3 and J48 a suitable prediction approach. Unlike ID3, J48 uses gain ratio as the root node selection metric, which is not biased towards input attributes with many values. We therefore use J48 throughout our analysis.

We discuss progressively sophisticated methods for predicting participants' decisions. After discussing naive solutions, we first present a cross-validated tree learning solution that results in a single "smart default" setting that is the same for everyone. Subsequently, we discuss three different procedures that create a number of "smart profiles" by clustering the participants and creating a separate cross-validated tree for each cluster. For each procedure, we try various numbers of clusters. Accuracies of the resulting solutions are reported in Table 5.

Naive Prediction Methods

We start with naive or "information-less" predictions. Our dataset contains 793 'yes'es and 2007 'no's. Therefore, predicting 'yes' for every scenario gives us a 28.33% prediction accuracy, while making a 'no' prediction gives us an accuracy of 71.67%. In other words, if we disallow all information collection by default, users will on average be happy with this default for 71.67% of the settings.

Overall Prediction

We next create a "smart default" by predicting the allow/reject decision with the scenario parameters using J48 with Weka's [11] default settings. The resulting tree (Figure 2) has an accuracy of 73.10%. The confusion matrix (Table 6) shows that this model results in overly conservative settings; only 208 'yes'es are predicted.

Figure 2 shows that this model predicts 'no' for every recipient (**who**) except 'Own device'. For this value, the default setting depends on **what** is being collected (see Table 7). For some levels of **what**, there is a further drill down based on **where**, **persistence** and **reason**.

Table 6: Confusion matrix for the overall prediction

Observed	Prediction		Total
	Yes	No	
Yes	124 (TP)	669 (FN)	793
No	84 (FP)	1923 (TN)	2007
Total	208	2592	2800

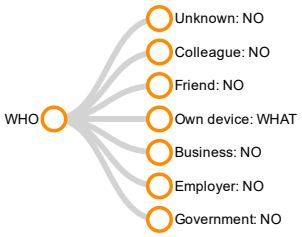

Figure 2: The Overall Prediction decision tree. Further drill down for **who** = 'Own device' is provided in Table 7

We can use this tree to create a "smart default" setting; in that case, users would on average be content with 73.10% of these settings—a 2% improvement over the naive "no to everything" default setting.

Given that people differ substantially in their privacy preferences, it is not unsurprising that this "one size fits all" default setting is not very accurate. A better solution would cluster participants by their privacy preferences, and then fit a separate tree for each cluster. These trees could then be used to create "smart profiles" that new users may choose from. Subsequent sections discuss several ways of creating such profiles.

Attitude-Based Clustering

Our first "smart profile" solution uses the attitudes (comfort, risk, appropriateness) participants expressed for each scenario on a 7-point scale. We averaged the values per attitude across each participant's 14 answers, and ran k-means clustering on that data with 2, 3, 4 and 5 clusters. We then added participants' cluster assignments to our original dataset, and ran the J48 decision tree learner on the dataset with the additional **cluster** attribute. Accuracies of the resulting solutions are reported in Table 5 under "attitude-based clustering".

All of the resulting trees had **cluster** as the root node. This indicates that this parameter is a very effective parameter for predicting users' decisions. This also allows us to split the trees at the root node, and create separate default settings for each cluster.

The 2-cluster solution (Figure 3) has a 75.28% accuracy — a 3.0% improvement over the "smart default". This solution results in one profile with 'no' for everything, while for the other profile the decision depends on the recipient (**who**). This profile allows any collection involving the user's 'Own device', and may allow collection by a 'Friend' or an 'Employer/School', depending on **what** is being collected.

Table 7: Drill down of the Overall Prediction tree for **who** = 'Own device'

What	Decision		
PhoneID	Yes		
PhoneID>identity	Yes		
Location	No		
Location>presence	Reason	Safety	Yes
		Commercial	Yes
		Social-related	No
		Convenience	No
		Health-related	Yes
		None	Yes
Voice	No		
Voice>gender	Where	Your place	No
		Someone else	No
		Semi-public	No
		Public	Yes
Voice> age	No		
Voice>identity	Yes		
Voice>presence	Yes		
Voice>mood	Yes		
Photo	No		
Photo>gender	No		
Photo>age	No		
Photo>identity	Yes		
Photo>presence	No		
Photo>mood	No		
Video	No		
Video>gender	No		
Video>age	No		
Video>presence	No		
Video>mood	Yes		
Video>looking at	Persistence	Once	Yes
		Continuous	No
Gaze	No		
Gaze>looking at	Reason	Safety	Yes
		Commercial	No
		Social-related	No
		Convenience	Yes
		Health-related	Yes
		None	Yes

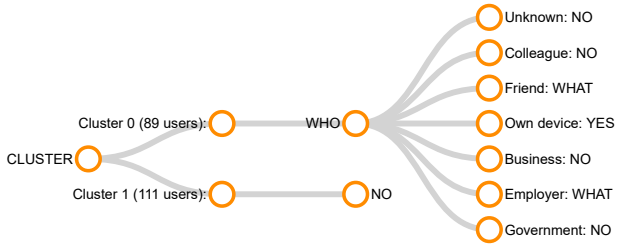

Figure 3: Attitude-based clustering: 2-cluster tree. Further drill down for **who** = 'Friend' or 'Employer/School' in Cluster 0 is hidden for space reasons.

The 3-cluster solution has a slightly lower accuracy of 75.17%, but is more parsimonious than the 2-cluster solution. There is one profile with 'no' for everything, one profile that allows collection by the user's 'Own device' only, and one profile that allows any collection except when the recipient is 'Unknown' or the 'Government'. The 4- and 5-cluster solutions have several clusters with the same sub-tree, and therefore reduce to a 3-cluster solution with 75.60% and 75.25% accuracy, respectively.

Fit-based clustering

Our fit-based clustering approach clusters participants without using any additional information. It instead uses the fit of the tree models to bootstrap the process of sorting participants into clusters. Like many bootstrapping methods, ours uses *random starts* and *iterative improvements* to find the optimal solution.

Random starts: We randomly divide particpants over N separate groups, and learn a tree for each group. This is repeated until a non-trivial starting solution (i.e., with distinctly different trees per cluster) is found.

Iterative improvements: Once each of the N groups has a unique decision tree, we evaluate for each participant which of the trees best represents their 14 decisions. If this is the tree of a different group, we switch the participant to this group. Once all participants are evaluated and put in the group of their best-fitting tree, the tree in each group is re-learned with the data of the new group members. This then prompts another round of evaluations, and this process continues until no further switches are performed.

Since this process is influenced by random chance, it is repeated in its entirety to find the optimal solution. Cross-validation is performed in the final step to prevent over-fitting. Accuracies of the 2- and 3-cluster solutions are reported in Table 5 under "fit-based clustering". We were not able to converge on a higher number of clusters.

The 2-cluster solution has a 77.99% accuracy—a 6.7% improvement over the "smart default". One profile has 'no' for everything, while the settings in the other profile depends on **who**: it allows any collection by the user's 'Own device', and may allow collection by a 'Friend's device' or an 'Employer', depending on **what** is collected.

The 3-cluster solution (Figure 4) has a 81.54% accuracy — an 11.5% improvement over the "smart default". We find one profile with 'no' for everything; one profile that may allow collection by the user's 'Own device', depending on **what** is being collected; and one profile that allows any collection except when the recipient (**who**) is 'Unknown', the 'Government', or a 'Colleague', with settings for the latter depending on the **reason**.

Agglomerative clustering

Our final method for finding "smart profiles" follows a hierarchical bottom-up (or agglomerative) approach. It

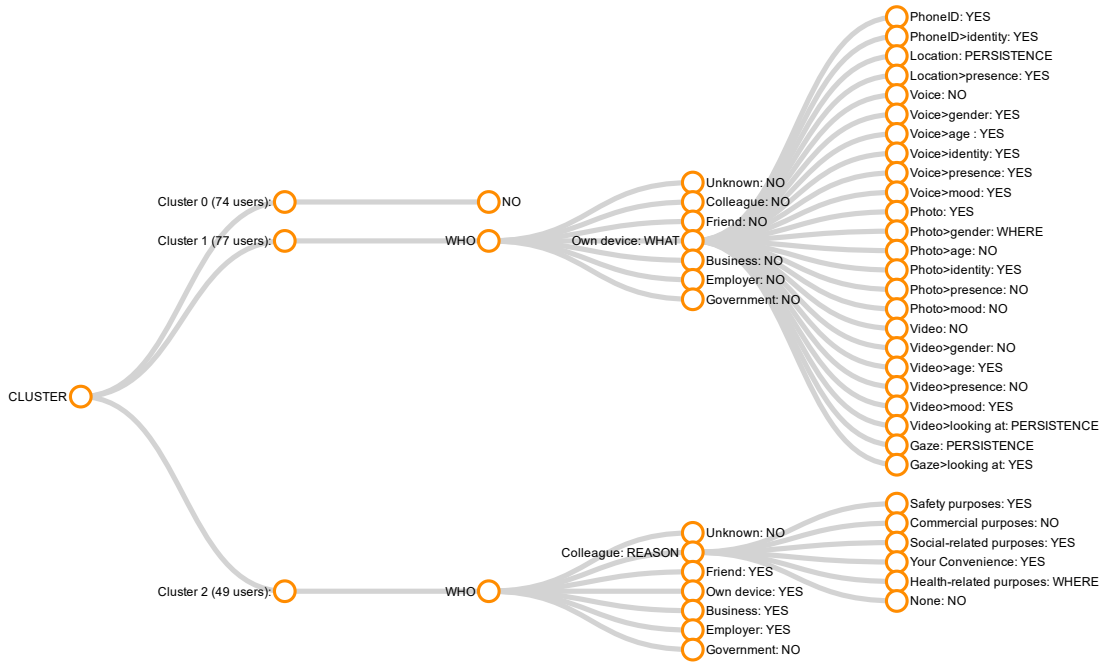

Figure 4: Fit-based clustering: 3-cluster tree. Further drill down is hidden for space reasons.

first fits a separate tree for each participant, and then iteratively merges them based on similarity. 156 of the initial 200 trees predict "no for everything" and 34 of them predict "yes for everything"—these are merged first. For every possible pair of the remaining 10 trees, the accuracy of the pair is compared with the mean accuracy the individual trees, and the pair with the smallest reduction in accuracy is merged. This process is repeated until we reach the predefined number of clusters.

We were able to reach a 5- and 4-cluster solution. The 3-cluster solution collapsed down into a 2-cluster solution with one profile of all 'yes'es and one profile of all 'no's (a somewhat trivial solution with a relatively bad fit). Accuracies of the 4- and 5-cluster (Table 5, "agglomerative clustering") are 78.13% and 78.27% respectively. For the 4-cluster solution, we find one profile with 'no' for everything, one profile with 'yes' for everything, one profile that depends on **who**, and another that depends on **what**. The latter two profiles drill down even further on specific values of **who** and **what**, respectively.

Discussion of Machine Learning Results

Figure 5 shows a comparison of the presented approaches. Compared to a naive default setting (all 'no'), a "smart default" makes a 2.0% improvement. The fit-based 2-cluster solution results in two "smart profiles" that make another 6.7% improvement over the "smart default", while the three "smart profiles" of the fit-based 3-cluster solution make an 11.5% improvement. If we let users choose the best option among these three profiles, they will on average be content with 81.54% of the settings. This rivals the accuracy of some of the "active tracking" machine learning approaches (cf. [24]).

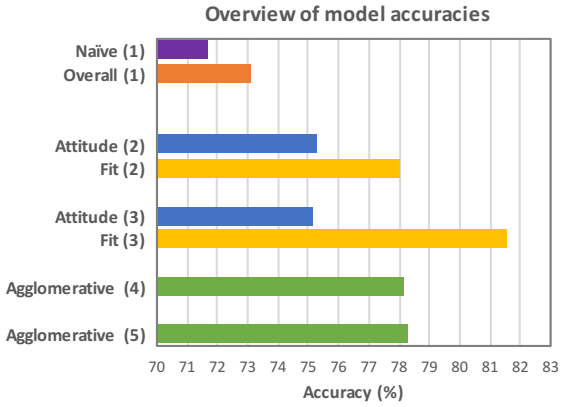

Figure 5: Accuracy of our clustering approaches

In line with our statistical results, the factor **who** seems to be the most prominent parameter, followed by **what**. In some cases the settings are more complex, depending on a combination of **who** and **what**. This is in line with the interaction effect observed in our statistical results.

Even our most accurate solution is not without fault, and its accuracy depends most on the **who** parameter. Specifically, the solution is most accurate for the user's own device, the device of a friend, and when the recipient is unknown. It is however less accurate when the recipient is a colleague, a nearby business, an employer, or the government. In these scenarios, more misclassifications tend to happen, so it would be useful to 'guide' users to specifically have a look at these default settings, should they opt to make any manual overrides.

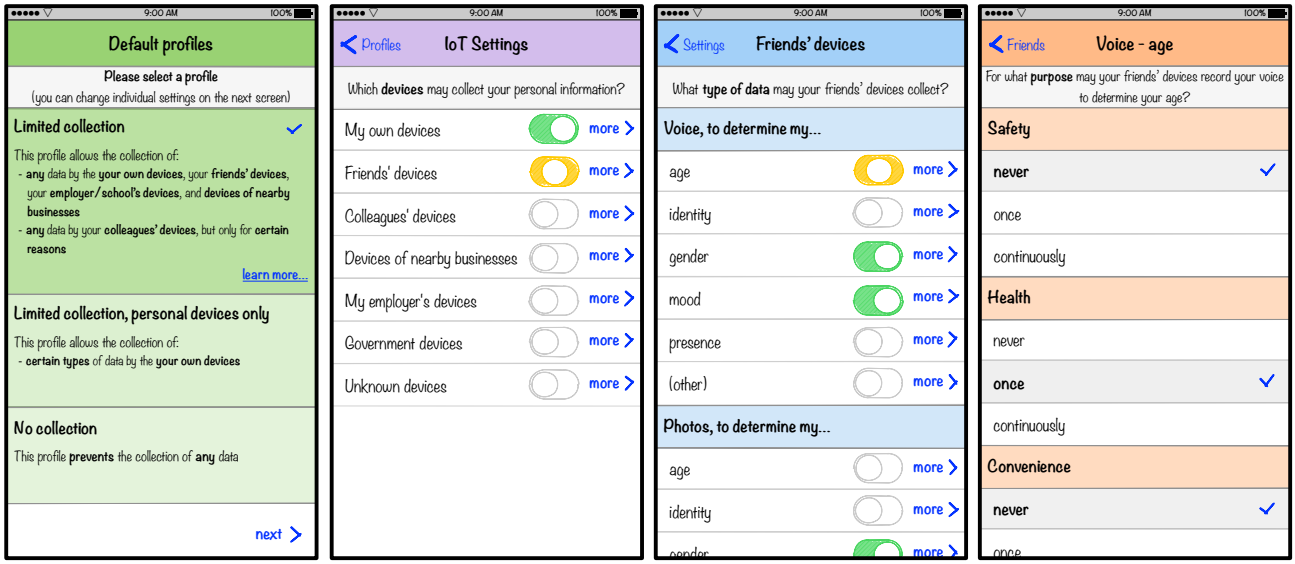

Figure 6: From Left, Screen 1 shows three default settings, Screen 2,3 and 4 shows layered interface

PRIVACY-SETTING PROTOTYPES

Designers of IoT privacy-setting interfaces face a difficult challenge. Since there currently exists no system for setting one's privacy preferences for public IoT scenarios, designers must rely on existing data such as the Lee and Kobsa [16] dataset to inform the design of these interfaces. Moreover, even for the simplified scenarios in this dataset, a privacy-setting interface will likely be complex, as it requires users to navigate settings for 7 types of recipients (**who**), 24 types of information (**what**), 4 different locations (**where**), 6 purposes (**reason**), and decide whether they want to allow the collection once or continuously (**persistence**). In this section we employ our data-driven design methodology to develop a prototype for an IoT privacy-setting interface based on the results of our statistical and machine learning analyses.

Manual Settings

The first challenge is to design an interface that users can navigate manually. Using the results of our statistical analyses, we design a "layered" settings interface: users can make a decision based on a single parameter only, and choose 'yes', 'no', or 'it depends' for each parameter value. If they choose 'it depends', they move to a next layer, where the decision for that parameter value is broken down by another parameter.

The manual interface is shown in Screens 2-4 of Figure 6. At the top layer of this interface should be the scenario parameter that is most influential in our dataset. Our statistical results inform us that this is the **who** parameter. Screen 2 shows how users can allow/reject data collection for each of the 7 types of recipients. Users can choose "more", which brings them to the second-most important scenario parameter, i.e. the **what** parameter. Screen 3 shows the data type options for when the user clicks on "more" for "Friends' devices". We have conveniently grouped the options by collection medium. Users can turn the collection of various data types by their friends' devices on or off. If only some types of data are allowed, the toggle at the higher level gets a yellow color and turns to a middle option, indicating that it is not completely 'on' (see "Friends' devices" in Screen 2).

Screen 4 shows how users can drill down even further to specify **reasons** for which collection is allowed, and the allowed **persistence** (we combined these two parameters in a single screen to reduce the "depth" of our interface). Since **reason** and **persistence** explain relatively little variance in behavioral intention, we expect that only a few users will go this deep into the interface for a small number of their settings. We leave out **where** altogether, because our statistical results deemed this parameter to be non-significant.

Smart Default Setting

The next challenge is to decide on a default setting, so that users only have to make minimal adjustments to their settings. We can use a simple "yes to everything" or "no to everything" default, but these are on average only accurate 28.33% and 71.67% of the time, respectively.

Using the results from our Overall Prediction (see Figure 2), we can create a "smart default" setting that is 73.10% accurate on average. In this version, the IoT settings for all devices are set to 'off', except for 'My own device', which will be set to the middle option. Table 7 shows the default settings at deeper levels. As this default setting is on average only 73.10% accurate, we expect users to still change some of their settings. They can do this by navigating the manual settings interface.

Smart Profiles

To improve the accuracy of the default setting, we can instead build two "smart profiles", and allow the user to

choose among them. Using the 3-cluster solution of the fit-based approach (see Figure 4), we can attain an accuracy of 81.54%. Screen 1 in Figure 6 shows a selection screen where the user can choose between these profiles. The "Limited collection" profile allows the collection of any information by the user's own devices, their friends' devices, their employer/school's devices, and devices of nearby businesses. Devices of colleagues are only allowed to collect information for certain reasons. The "Limited collection, personal devices only" profile only allows the collection of certain types of information by the user's own devices. The "No collection" profile does not allow any data collection to take place by default.

Once the user chooses a profile, they will move to the manual settings interface (Screens 2–4), where they can further change some of their settings.

CONCLUSION

The motivation behind our research was the information and choice overload associated with the plethora of choices that users might face while setting their privacy settings in an IoT environment. We have made use of statistical analyses and machine learning algorithms to provide a data-driven design for an IoT privacy-setting interface. We summarize this procedure as follows:

- Using statistical analysis, uncover the relative importance of the parameters that influence users' privacy decisions. Develop a "layered interface" in which these parameters are presented in decreasing order of importance.
- Using a tree-learning algorithm, create a decision tree that best predicts participants' choices based on the parameters. Use this tree to create a "smart default" setting.
- Using a combination of clustering and tree-learning algorithms, create a set of N decision trees that best predict participants' choices. Use the trees to create N "smart profiles".
- Develop a prototype for an IoT privacy-setting interface that integrates the layered interface with the smart default or the smart profiles.

We demonstrated this procedure by applying it to a dataset collected by Lee and Kobsa [16]. In the process, we made a number of interesting observations.

The statistical and machine learning results both indicated that recipient of the information (**who**) is the most significant parameter in users' decision to allow or reject IoT-based information collection. This parameter therefore features at the forefront in our layered settings interface, and plays an important role in our smart profiles.

The **what** parameter was the second-most important decision parameter, and interacted significantly with the **who** parameter. This parameter therefore features at the second level of our settings interface, and further qualifies some of the settings in our smart profiles.

Our layered interface allows a further drill-down to the **reason** and **persistence** parameters, but given the relatively lesser importance of these parameters, we expect few users to engage with the interface at this level. Moreover, the **where** parameter was not significant, so we left it out of the interface.

While a naive ('no' to all) default setting in our interface would have provided an accuracy of 71.67%, it would not have allowed users to reap the potential benefits associated with IoT data collection without changing the default setting. Our Overall Prediction procedure resulted in a smart default setting that was a bit more permissive, and increased the accuracy by 2%.

The fit-based clustering approach, which iteratively clusters users and fits an optimal tree in each cluster, provided the best solution. This resulted in an interface where users can choose from 3 profiles, which increases the accuracy by another 11.5%.

Our analysis allowed us to use *data-driven design* to bootstrap the development of a privacy-setting interface, but a future user experiment could investigate whether users are comfortable with the layered interface, and whether they prefer a single "smart default" setting or a choice among "smart profiles".

The scenario-based method presented in this paper is particularly suited for novel domains where few real interaction exist. We note, though, that this novelty may hamper our approach: users' decisions are inherently limited by the knowledge they have about IoT. Lee and Kobsa [16] made sure to educate users about the presented scenarios, hence their data is arguably better in this regard than data from "live" systems. However, as the adaptation of IoT becomes more widespread, the mindset and knowledge regarding such technologies—and thus their privacy preferences—might change. Our "smart profiles" may thus eventually have to be updated in future work, but for now, our current profiles can at least help users make make better privacy decisions in their initial stages of usage.

Future work could also apply the proposed procedure to other privacy-setting domains. In using scenarios, the procedure avoids typical decision externalities such as default effects, framing effects, and decision-context effects that tend to obfuscate users' behaviors in more naturalistic studies. Moreover, the scenarios can inform the creation of privacy-setting interfaces for novel or currently non-existent technologies. As such we imagine that the procedure could be applied in new domains, such as household IoT ("smart home") privacy, drone privacy, and nano-tech privacy. In some of these domains, fully "adaptive" privacy mechanisms that use "active tracking" (cf. [13, 18]) are more suitable, while other domains could benefit from our static, profile-based approach.

REFERENCES

1. Acquisti, A., and Gross, R. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *International workshop on privacy enhancing technologies* (2006), Springer, pp. 36–58.
2. Ajzen, I., and Fishbein, M. Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin 84*, 5 (1977).
3. Baron, R. M., and Kenny, D. A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology 51*, 6 (1986), 1173.
4. Boyles, J. L., Smith, A., and Madden, M. Privacy and Data Management on Mobile Devices. Tech. rep., Pew Internet & American Life Project, 2012.
5. Chow, R., Egelman, S., Kannavara, R., Lee, H., Misra, S., and Wang, E. HCI in Business: A Collaboration with Academia in IoT Privacy. In *HCI in Business*, F. F.-H. Nah and C.-H. Tan, Eds., no. 9191 in Lecture Notes in Computer Science. Springer International Publishing, 2015.
6. Dong, C., Jin, H., and Knijnenburg, B. P. Ppm: A privacy prediction model for online social networks. In *International Conference on Social Informatics* (2016), Springer, pp. 400–420.
7. Fang, L., and LeFevre, K. Privacy wizards for social networking sites. In *Proceedings of the 19th international conference on World wide web* (2010), ACM, pp. 351–360.
8. Good, N., Dhamija, R., Grossklags, J., Thaw, D., Aronowitz, S., Mulligan, D., and Konstan, J. Stopping Spyware at the Gate: A User Study of Privacy, Notice and Spyware. In *Proceedings of the 2005 Symposium on Usable Privacy and Security* (2005), ACM, pp. 43–52.
9. Gross, R., and Acquisti, A. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society* (2005), ACM, pp. 71–80.
10. Gubbi, J., Buyya, R., Marusic, S., and Palaniswami, M. Internet of things (iot): A vision, architectural elements, and future directions. *Future generation computer systems 29*, 7 (2013), 1645–1660.
11. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. The weka data mining software: an update. *ACM SIGKDD explorations newsletter 11*, 1 (2009), 10–18.
12. Jensen, C., and Potts, C. Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices. In *2004 Conference on Human Factors in Computing Systems* (2004), pp. 471–478.
13. Knijnenburg, B. P. *A user-tailored approach to privacy decision support*. Ph.D. Thesis, University of California, Irvine, Irvine, CA, 2015.
14. Knijnenburg, B. P., Kobsa, A., and Jin, H. Dimensionality of information disclosure behavior. *International Journal of Human-Computer Studies 71*, 12 (2013), 1144–1162.
15. Lederer, S., Mankoff, J., and Dey, A. K. Who wants to know what when? privacy preference determinants in ubiquitous computing. In *CHI'03 extended abstracts on Human factors in computing systems* (2003), ACM, pp. 724–725.
16. Lee, H., and Kobsa, A. Understanding user privacy in internet of things environments. *Internet of Things (WF-IoT)* (2016).
17. Li, Y., Kobsa, A., Knijnenburg, B. P., and Nguyen, M. C. Cross-cultural privacy prediction. *Proceedings on Privacy Enhancing Technologies 2* (2017), 93–112.
18. Liu, B., Andersen, M. S., Schaub, F., Almuhimedi, H., Zhang, S. A., Sadeh, N., Agarwal, Y., and Acquisti, A. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions. In *Proceedings of the 2016 Symposium on Usable Privacy and Security* (2016).
19. Madejski, M., Johnson, M., and Bellovin, S. M. A study of privacy settings errors in an online social network. In *IEEE International Conference on Pervasive Computing and Communications Workshops* (2012), IEEE, pp. 340–345.
20. Olson, J. S., Grudin, J., and Horvitz, E. A study of preferences for sharing and privacy. In *CHI'05 extended abstracts on Human factors in computing systems* (2005), ACM, pp. 1985–1988.
21. Pallapa, G., Das, S. K., Di Francesco, M., and Aura, T. Adaptive and context-aware privacy preservation exploiting user interactions in smart environments. *Pervasive and Mobile Computing 12* (2014), 232–243.
22. Raber, F., Luca, A. D., and Graus, M. Privacy wedges: Area-based audience selection for social network posts. In *Proceedings of the 2016 Symposium on Usable Privacy and Security* (2016).
23. Ravichandran, R., Benisch, M., Kelley, P. G., and Sadeh, N. M. Capturing social networking privacy preferences. In *Proceedings of the 2009 Symposium on Usable Privacy and Security* (2009), Springer, pp. 1–18.

24. Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., and Rao, J. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing 13*, 6 (2009), 401–412.

25. Sandhu, R. S., and Samarati, P. Access control: principle and practice. *IEEE Communications Magazine 32*, 9 (1994), 40–48.

26. Smith, N. C., Goldstein, D. G., and Johnson, E. J. Choice Without Awareness: Ethical and Policy Implications of Defaults. *Journal of Public Policy & Marketing 32*, 2 (2013), 159–172.

27. Watson, J., Besmer, A., and Lipford, H. R. +Your circles: sharing behavior on Google+. In *Proceedings of the 8th Symposium on Usable Privacy and Security* (2012), ACM, pp. 12:1–12:10.

28. Williams, M., Nurse, J. R., and Creese, S. The perfect storm: The privacy paradox and the internet-of-things. In *Availability, Reliability and Security (ARES), 2016 11th International Conference on* (2016), IEEE, pp. 644–652.

29. Wisniewski, P. J., Knijnenburg, B. P., and Lipford, H. R. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of Human-Computer Studies 98* (2017), 95–108.