ISR Institute for Software Research
University of California, Irvine

Towards Ubiquitous Privacy Decision Support: Machine Prediction of Privacy Decisions in IoT

Hosub Lee
University of California, Irvine
hosubl@uci.edu

Alfred Kobsa
University of California, Irvine
kobsa@uci.edu

October 2018
ISR Technical Report # UCI-ISR-18-6

Institute for Software Research
ICS2 221
University of California, Irvine
Irvine, CA 92697-3455
isr.uci.edu

isr.uci.edu/publications

Towards Ubiquitous Privacy Decision Support: Machine Prediction of Privacy Decisions in IoT[*]

Hosub Lee and Alfred Kobsa
Institute for Software Research &
Donald Bren School of Information and Computer Sciences
University of California, Irvine
{hosubl,kobsa}@uci.edu

Abstract

We present a mechanism to predict privacy decisions of users in Internet of Things (IoT) environments, through data mining and machine learning techniques. To construct predictive models, we tested several different machine learning models, combinations of features, and model training strategies on human behavioral data collected from an experience-sampling study. Experimental results showed that a machine learning model called linear model and deep neural networks (LMDNN) outperforms conventional methods for predicting users' privacy decisions for various IoT services. We also found that a feature vector, composed of both contextual parameters and privacy segment information, provides LMDNN models with the best predictive performance. Lastly, we proposed a novel approach called one-size-fits-segment modeling, which provides a common predictive model to a segment of users who share a similar notion of privacy. We confirmed that one-size-fits-segment modeling outperforms previous approaches, namely individual and one-size-fits-all modeling. From a user perspective, our prediction mechanism takes contextual factors embedded in IoT services into account and only utilizes a small amount of information polled from the users. It is therefore less burdensome and privacy-invasive than the other mechanisms. We also discuss practical implications for building predictive models that make privacy decisions on behalf of users in IoT.

Keywords— Privacy decision prediction, linear model and deep neural networks, privacy segmentation, K-modes clustering, Internet of Things

1 Introduction

Smartphone apps and websites increasingly ask users to make privacy decisions about personal information disclosure, e.g., to grant or deny an app permission to access their

[*] ISR Technical Report #UCI-ISR-18-6, October 2018

location or their phone book. However, previous research indicates that people are often unable to make these decisions due to limits in their available time, motivation, and abilities to fully understand the tradeoff between utility and privacy in the disclosure of personal information. This problem will continue to grow in ubiquitous computing environments like the Internet of Things (IoT), as an array of sensors around the user unobtrusively collects his/her personal information [3]. Clearly, rich personal information helps IoT systems better understand users, thus providing better-tailored services to them. At the same time, however, this leads to considerable increase in privacy concerns that may lead users to stop using the service [10, 39, 33]. Therefore, providing IoT services with minimized privacy risks is crucial for both protecting users' privacy and keeping IoT ecosystems sustainable. One possible way to achieve this objective is to assist users with making better privacy decisions, by predicting decisions based on their historical decision-making behavior and recommending privacy settings accordingly (i.e., privacy decision support). Most previous research on privacy decision support has focused on personal information disclosure in online or mobile social network services. Researchers employed supervised machine learning approaches that build predictive models by utilizing users' past behavior as training data, and then utilized the trained models to recommend the most appropriate privacy decisions in the current situation [35, 12, 36, 37, 40, 6]. They also verified that privacy decision support systems alleviate users' cognitive burden, thereby allowing them to make their preferred decisions more easily. This kind of technology will become more crucial in IoT environments, not only because users will need to make privacy decisions more frequently, but also because there are no or only very limited user interfaces available for users to state their preferences or decisions to the services (e.g., there are no standard keyboards and displays).

In this paper, we proposed a novel machine learning mechanism for predicting privacy decisions of users in IoT environments. The aim of this mechanism is to correctly predict users' decisions whether or not to allow the given personal information monitoring based on both the user's current context and personal attitudes on privacy. We tested the proposed mechanism on a privacy-related behavioral dataset collected from 172 users who were presented with descriptions of personal information tracking scenarios relating to their physical location on a university campus [24]. These scenarios are defined by five different contextual parameters such as place (*where*), type of collected information (*what*), entity (*who*), purpose (*reason*), and frequency (*persistence*) of the monitoring. The dataset also contains each user's privacy decision behavior on each scenario in terms of five reaction parameters such as willingness to be notified (*_notification*), willingness to allow (*_permission*), and subjective evaluations of comfort, risk, and appropriateness (*_comfort*, *_risk*, and *_appropriateness*) of the monitoring. We treated the five contextual parameters as basic features which collaboratively represent the current context in which information monitoring is performed by IoT devices. Additionally, we assigned each user to a specific privacy perception segment based on a small portion of their data (i.e., prior privacy decisions). We then used this privacy segment information as an additional feature. We used the reaction parameter *_permission* as target value (or label), since it best reflects users' substantive privacy decisions in IoT environments (namely, allow or reject the monitoring). The dataset contains 6,618 data points, and each data point (row) represents a user's privacy decisions in a specific IoT scenario.

First, we utilized a state-of-the-art machine learning model, called linear model and deep neural networks (LMDNN, [8]), to make privacy decisions on behalf of users. LMDNN, which is also known as Wide & Deep Learning, jointly trains wide linear models and deep neural networks. By doing so, it can take the benefits of memorization (linear models) and generalization (neural networks) at the same time. LMDNN showed a remarkable performance on binary classification problems with sparse input features [8, 38]. Because the dataset collected in [24] is also composed of categorical data with many possible feature values (e.g., 24 values for the contextual parameter *what*), we decided to use LMDNN to build predictive models for privacy decision support in IoT. We also selected machine learning models that have been widely used in the literature (e.g., decision trees) and compared them with LMDNN in terms of predictive performance on the dataset. Experimental results indicated that LMDNN outperforms all the conventional models.

Next, we explored the most suitable combination of features for building LMDNN models with a reasonable predictive performance. We chose the five contextual parameters as basic features because these parameters are known to be related to users' privacy decisions in IoT [9, 23, 24]. In addition, we considered each user's privacy segment information as an additional feature. This is because previous research indicates that privacy segment information is helpful for machine learning models to better predict users' privacy decisions [27, 28]. We applied an unsupervised data clustering algorithm, K-modes clustering, on a subset of users' privacy decisions, in order to segment the users by their perceptions of privacy (i.e., privacy segmentation). Therefore, we tested the following feature combinations: (1) contextual parameters only (basic features), (2) contextual parameters with interactions between them, and (3) contextual parameters with interactions between them and privacy segment information. Experimental results showed that the feature combination (3) gives LMDNN models with the highest predictive performance. It also means that both interactions between contextual factors (e.g., *who* by *what*) and privacy segment information are useful for LMDNN models to predict privacy decisions of the users.

Lastly, we investigated different approaches for training machine learning models. There exist two traditional approaches in the literature: individual and one-size-fits-all modeling. Individual modeling is a process of building a user-specific predictive model based on each user's data only. This approach is known to be effective for modeling each user's unique characteristics (e.g., habits and personality). A model with a reasonable predictive performance, however, typically requires a considerable amount of training data from each individual user. In contrast, one-size-fits-all modeling utilizes multiple users' data as a single training data and constructs a universal model for all of them (including new users). This approach enables predictive models to make general predictions, therefore it can be useful for new users who did not provide data to the system yet, but want to get recommendations immediately. However, prediction results may not be personalized to each individual user. We present another approach called one-size-fits-segment modeling, a variant of one-size-fits-all, taking privacy segment information into account in building predictive models. The basic idea is to serve each user with a machine learning model trained by data collected from others who share the same notion of privacy with this user. We divided the dataset based on the results of privacy segmentation, then trained an LMDNN model for each segment of the users. By using both contextual and privacy segment information as input features, we compared the predictive performance of individual, one-size-fits-all, and one-size-fits-segment modeling. Experimental results confirmed that the proposed approach performs the best in the dataset. Final LMDNN models trained via one-size-fits-segment modeling showed an average area under curve (AUC) of 0.6782 across all users. We noticed that one-size-fits-segment modeling performs much better than individual modeling for about 80% of the users. However, it does not work well for about 20% of the users who have highly accurate individual models ($AUC > 0.7$).

To sum up, our proposed prediction mechanism not only showed a reasonable performance for most users, but also can cause less burden and privacy risks to users since it mainly utilizes non-personal contextual information which can often be automatically collected from the IoT environment, and only prompts each user for a small amount of privacy decisions (answers to five reaction parameters about a single scenario) to determine his/her privacy segment. We also presented some practical implications for designing and developing machine learning-based privacy decision support systems for IoT.

In summary, our work makes the following contributions to the field of privacy decision support:

1. We adopted a state-of-the-art machine learning model called linear model and deep neural networks (LMDNN) to make privacy decisions on behalf of IoT users, and verified that LMDNN outperforms conventional methods that have been widely used in the literature;
2. We investigated the best approach for training machine learning models in terms of input feature and training strategy, and verified that the proposed approach called one-size-fits-segment modeling outperforms preexisting methods such as individual and one-size-fits-all modeling;
3. We reported some practical implications in predicting users' privacy decisions on personal information monitoring in IoT.

2 Related Work

In this section, we first present a literature review of privacy decision prediction based on machine learning techniques. As mentioned above, a considerable research effort has been made towards developing intelligent agents which infer and recommend users' privacy decisions on diverse personal information disclosure scenarios. Researchers have also claimed that this kind of technology could help users make a correct decision, thereby minimizing potential privacy risks. In addition, we also summarized previous research about privacy segmentation. It is understood that privacy segmentation provides useful information for understanding and predicting people's privacy decision-making.

2.1 Prediction of Privacy Decision-making

Much research on the prediction of privacy decision-making focused on online or mobile social network services (SNS). By using (semi-) supervised machine learning techniques, researchers aimed to accurately predict SNS users' sharing policies for their own contents (e.g., to allow or disallow Facebook friend John to see my photos). This is because vendor-provided privacy setting mechanisms, which ask users to manually specify their preferences, have been proven ineffective for protecting user privacy [35, 40, 6].

Fang et al. proposed a framework named Privacy Wizard that infers and recommends each user's access control policies for his/her personal information on Facebook [12]. Each policy specifies who can access specific personal information. The authors adopted a supervised machine learning approach to learn each user's policies by asking him/her a number of questions (e.g., would you like to share your birthday with Facebook friend John?). Users' answers were considered intended policies (labels) for their friends. Regarding input features for machine learning models, the authors utilized both demographic information and community membership which characterized each user. Most importantly, the framework asked users about policies that machine learning models are most uncertain about (namely, active learning with uncertainty sampling). By selectively asking users to label the most informative data first, active learning can effectively reduce manual labeling efforts. In the experiments with 45 Facebook users, Privacy Wizard showed 90% accuracy in predicting individual's privacy policies with a small amount of labeled training data (25 out of 200 friends with privileges). The authors utilized a decision tree as a machine learning model.

Shehab et al. proposed a framework called Policy Manager that predicts users' sharing policies on SNS [36]. Like [12], Policy Manager was designed to infer binary access control policies on a specific personal object posted on each user's SNS account. The authors used both user profiles (that included, e.g., gender) and social network structure (e.g., closeness among users) as input features, and asked users to manually determine policies for a subset of their friends (labeling of training data). They tried nine different machine learning algorithms for each user, and chose the best algorithm showing the highest cross-validation accuracy on his/her data. In addition, Policy Manager selected models (i.e., classifiers) trained by other users, based on their accuracy in predicting a target user's training data. It then utilized these classifiers with the target user's own classifier to produce a final classification (i.e., classifier fusion via group voting). The authors tested their framework with 200 Last.FM users, and confirmed that the proposed classifier fusion approach was effective in improving predictive performance compared to using an individual classifier only: 83% and 70% best accuracy when using an alternating decision tree (ADTree) machine learning model, respectively. The authors extended their work by applying an active learning paradigm into the semi-supervised learning framework [37]. Similar to [12], they aimed at minimizing user burden in labeling training data by selecting the most informative data points to label. Sinha et al. also developed automated tools to assist users in correctly configuring privacy policies for text-based content on Facebook [40]. The authors utilized diverse features such as the text of posts, time of creation, n-grams, the previous policy, and attachments to predict future policies (e.g., visible to only me) for Facebook posts. They also adopted a supervised learning approach based on the MaxEnt algorithm. Through an online survey study with real Facebook users, they found out that the system could predict policies with a maximum accuracy of 81%, leading to a 14% increase in accuracy compared to when users used Facebook's privacy setting mechanism for a new post.

Regarding mobile SNS, Sadeh et al. proposed a mobile social networking application, PeopleFinder, which recommends optimized location privacy policies to users [35]. To this end, the authors adopted random forests, an ensemble supervised learning method, to build a classifier predicting sharing policies for each user's current location based on his/her previous decisions. The authors proved that these machine-generated policies have better accuracy than the user-defined policies: 91% and 79% success rate in matching users' actual behavior, respectively. Recently, Bilogrevic et al. proposed a personal information sharing platform named SPISM that semi-automatically determines whether to disclose users' personal information on SNS and at what level of granularity [6]. Like other previous works, the authors used a supervised learning approach to predict SNS users' privacy decision-making. For each user, SPISM constructed a multi-level classifier based on nave Bayes or support vector machine (SVM) by using his/her past behavior as training data. Training data are composed of diverse personal and contextual factors, such as the identity of requester, type of information requested, user location, co-presence of others, and time (features) and each user's past privacy decision (label). Like [12, 37], SPISM also adopted an active learning paradigm to minimize users' labeling efforts. SPISM made decisions automatically whenever the confidence (probability) in the classification result was high enough; otherwise it explicitly asked users' decisions then added them to the pre-existing training data. With the updated training data, SPISM continuously learns and adapts to users' privacy behaviors. Therefore, it will require less and less user input over time. A user study with 70 participants indicated that SPISM outperforms user-defined policies; it showed a median prediction accuracy of 72% when each user provided 40 manual decisions. Even if the authors focused on building a personalized machine learning model for each user (i.e., individual modeling), they also assessed potentials of one-size-fits-all modeling. A universal model trained by all users' data showed a reasonable performance with a median accuracy of 67%. The authors claimed that one-size-fits-all modeling could be suitable for building an initial predictive model that produces default privacy settings for the new user.

All these works not only confirm the necessity of decision support systems for protecting user privacy in social network services, but also provide practical guidelines for learning people's privacy behavior which often evolves over time. Specifically, most previous works were based on supervised machine learning algorithms with an individual modeling approach. That is, researchers proposed ways to build user-specific predictive models considering each user's unique behavioral characteristics. To reduce the user burden in providing sample privacy decisions, the researchers let users respond to select privacy-invasive scenarios or situations (i.e., active learning).

2.2 Privacy Segmentation

Researchers have also investigated methodologies to segment users into several categories in terms of their privacy attitudes and behaviors. The most commonly cited methodology is Westin's privacy segmentation model [21]. Westin had conducted several surveys about privacy issues in various domains such as e-commerce, national identification systems, and e-health. To effectively summarize the survey results, Westin developed an indexing scheme that categorizes survey participants into three categories: Privacy Fundamentalists, Privacy Pragmatists, and Privacy Unconcerned. Westin treated participants' responses to several pre-defined statements (scenarios) as criteria to derive these three categories (e.g., Privacy Fundamentalists are respondents who agreed with the first statement and disagreed with the second and third statement).

Privacy segmentation has been studied in diverse contexts. Lin et al. proposed an unsupervised data clustering approach for categorizing smartphone users into distinctive groups based on their privacy preferences regarding mobile app permission (e.g., grant or deny permission to an app to access personal information) [27]. The authors utilized an agglomerative hierarchical clustering algorithm (Ward's method) on about 21,000 privacy preferences collected from 725 Android users. Each preference represents each user's willingness to grant permission to a given app for a specific purpose (i.e., app-permission-purpose triple). The authors identified four privacy profiles from this cluster analysis. They also presented default privacy settings to each user based on his/her privacy profile. This was intended to help users better control their privacy when confronted with numerous permission requests on Android platforms. In a follow-up study, the authors utilized users' privacy profile information as one of the input features for machine learning models to predict Android users' permission settings [28].

Lankton et al. clustered SNS users into four categories based on their privacy management strategies [22]. They surveyed college students' behavior on Facebook, including the use of privacy settings, degree of content disclosure, and variety and size of friend lists. The authors then conducted a two-stage cluster analysis on this dataset. Like [27], the authors first performed hierarchical clustering to determine both the correct number of clusters and initial cluster centroids. They found that a four-cluster solution is optimal. Next, they conducted non-hierarchical (K-means) cluster analysis on the dataset, using the pre-determined cluster centroids as a starting point. After statistically comparing survey responses in each cluster, the authors confirmed that the resulting clusters are distinctive enough regarding the degree of the users' privacy concerns.

Most recently, the market research firm Forrester published a report suggesting that consumers can be divided into four privacy categories: Data-Savvy Digitals, Reckless Rebels, Nervous Nellies, and Skeptical Protectionists [19]. This finding is based on large-scale online survey studies designed to capture people's behavioral reactions toward personal data collection and use by Internet companies. About 34% of the study participants were categorized as those who are not willing to share their information (Nervous Nellies and Skeptical Protectionists) since they are skeptical about corporate privacy practices.

Even though the number and type of privacy segments vary somewhat across these works, most researchers came to the same common conclusion about privacy segmentation: it is practically feasible to identify distinctive privacy segments by collecting and analyzing human behavioral data. Furthermore, privacy segment information can be used as an informative feature for understanding and predicting people's future privacy behavior because it represents users' overall perception of privacy [27, 28].

3 Dataset of Privacy Decisions

For this study we used the dataset collected in our previous work [24], which is arguably the first comprehensive collection of privacy decisions in IoT environments. We gathered people's privacy preferences (decisions) about diverse privacy-invasive scenarios in simulated IoT environments, through the experience sampling method (ESM). We used Google Glass for presenting the IoT scenarios to study participants in order to let them perceive the scenarios as realistically as possible. Specifically, we developed a Google Glass app to dynamically display scenarios based on participants' location. Participants were asked to walk around a university campus wearing Google Glass. As participants moved towards one of 130 selected locations on campus, the app presented the scenario pertaining to this location. Participants then answered several questions on their preferred privacy protection in the given scenario. This immersive spatial setup seems more suitable to gather accurate privacy behaviors from participants than a traditional online survey system, since the setup situates them in scenarios and is therefore likely to better capture the situatedness of privacy decisions [13, 31]. In addition, location has been found to be a particularly critical component in understanding people's privacy decision-making [35, 5, 6].

In order to formalize users' privacy decision-making in IoT, we defined several parameters representing both contextual characteristics of IoT scenarios (contextual parameters) and possible user reactions (reaction parameters).

In our earlier interview and online survey studies [9, 23], we had already identified the five contextual parameters that have the most influence on people's reactions toward privacy (see Table A1). These five parameters define the place where the monitoring occurs (parameter *where*), the type of information being monitored (*what*), the entity that is monitoring (*who*), the reason for monitoring (*reason*), and the frequency of the monitoring (*persistence*). Each IoT scenario can be described by an expression that includes every contextual parameter together with its respective parameter value for this scenario. We produced 130 IoT scenarios specifically related to known geographical locations on the campus.

We also identified reaction parameters that serve as proxies of people's notion of privacy, namely the desire to be notified about (parameter *_notification*) and the willingness to allow (*_permission*) the monitoring. In addition, we also found it important to measure people's opinion on each monitoring activity in terms of comfort, risk, and appropriateness (parameters *_comfort*, *_risk*, *_appropriateness*). Table A2 shows the type of reaction parameters, together with their values which are all categorical or ordinal. A single user response about a specific IoT scenario (i.e., one row in a dataset) is composed of the following attributes: participant ID, scenario ID, five contextual parameter values, and the five user-provided reaction parameter values.

We gathered 172 participants in total over a period of three months: 106 males and 65 females (one person did not disclose his/her gender), with the majority (82%) being 18-25. Because we recruited the participants on campus, most of them have some university affiliation (about 2/3 undergraduates and 1/3 graduates). Participants answered 39 scenario descriptions on average. The collected dataset contains a total of 33,090 privacy decisions for 6,618 IoT scenarios. The dataset will be referred to as IoTP in this paper.

4 Privacy Decision Prediction

By using the IoTP dataset, we investigated mechanisms to learn and predict users' privacy decision-making in IoT environments. Specifically, we aimed to predict the value of the reaction parameter *_permission* (R_2) based on the current context and privacy segment of the user. Even though this parameter denotes four possible privacy decisions (see Table A2), we focused on binary decisions by converting R_2=1, 2 into R_2=1 (*allow*) and R_2=3, 4 into R_2=0 (*reject*). Therefore, prediction for this parameter can be formalized into binary classification problems (*allow* or *reject* the monitoring).

In this vein, we conducted a series of machine learning experiments, varying the models (algorithms), features, and training strategies. First, we tested multiple machine learning models to find the most suitable for making predictions about privacy decisions in IoT. We first tried LMDNN since it is known to be very effective for processing categorical data with high sparsity. We then compared the performance of LMDNN with several machine learning models that have been extensively used in earlier research. Thereafter we also assessed the impact of input features, consisting of contextual and privacy segment information, on the predictive power of the machine learning model. This assessment allowed us to determine which features should be used for building classifiers. Last, and most importantly, we conducted a comparative evaluation of the predictive performance of well-known model training strategies, such as individual and one-size-fits-all modeling, and our proposed approach called one-size-fits-segment modeling. The results informed us of practical implications for developing a privacy decision support system for IoT environments.

4.1 Machine Learning Model

In this sub-section, we explained in detail why we chose the LMDNN machine learning model for realizing privacy decision support in IoT. We also presented experimental results showing that LMDNN can provide the most reasonable predictive performance compared to conventional machine learning models used in the literature.

4.1.1 LMDNN

Linear model and deep neural networks (LMDNN), also known as Wide & Deep Learning, has been proposed by [8] to solve the problem of recommending apps on Google Play. Generalized linear models like logistic regression are widely used for large-scale regression and classification problems as they are simple, scalable, and interpretable. The models are often trained on binarized sparse input features with one-hot encoding. Memorization of diverse feature interactions can be efficiently achieved by feeding a wide set of cross-product feature transformations with a target value (or label) into the model. While linear models are effective for learning relationships between categorical features and a target value, they cannot generalize the relationships to identify feature-target patterns that do not exist in training data. In contrast, deep neural networks (DNN) can better generalize to the previously unseen patterns by using low-dimensional dense embedding vectors learned from the sparse input features (i.e., transforming a categorical feature value into a vector of continuous values). This means that DNN can make a reasonable prediction for new observations based on preexisting training data. At the same time, DNN also can over-generalize when the underlying feature-target matrix is too sparse (e.g., rare interactions between features and target value). LMDNN is a mixture of logistic regression (wide learning) and DNN (deep learning) to achieve both memorization and generalization in a single model. By jointly performing wide and deep learning, LMDNN complements the weakness of deep learning (i.e., over-generalization) by letting the wide learning take some cross-product feature transformations into account in producing final classifications. This Wide & Deep Learning paradigm shows a remarkable predictive performance on diverse classification problems [8, 7, 38].

Fig. 1 shows the LMDNN model structure we used in this study. The bottom (input) layer receives training data composed of categorical features (contextual parameters) together with a target value (binarized reaction parameter *_permission*). Next, the model can generate cross-product feature transformations and/or dense embedding vectors from the inputted data. As explained above, feature transformations and embedding vectors are used by wide and deep learning, respectively. For the deep learning, an 8-dimensional embedding vector is learned from each categorical feature. The model also combines all embedding vectors into a single dense embedding vector. The resulting concatenated embedding vector is then fed into three hidden layers with the ReLU activation function, and finally the logistic output unit (here, the sigmoid function). We configured 100, 50, and 100 units in consecutive hidden layers, respectively. We determined this network structure based on internal performance benchmarking on the IoTP dataset. In training LMDNN models, we followed the default mechanism (e.g., backpropagation, mini-batch stochastic optimization, AdaGrad regularization) described in [8]. We utilized a TensorFlow [1] implementation of

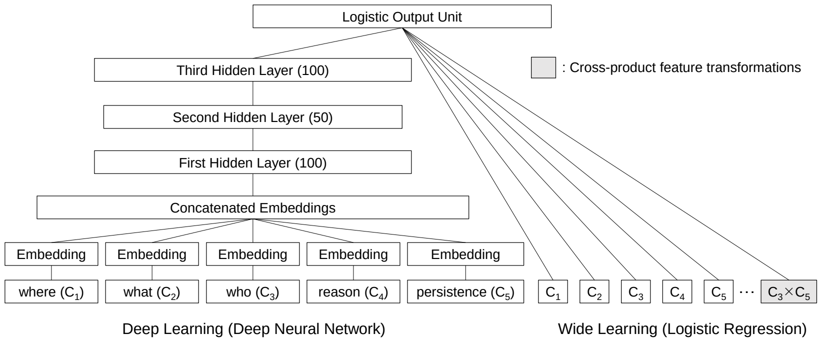

Figure 1: Architecture of the LMDNN model

LMDNN for all our experiments.

4.1.2 Predictive Performance Evaluation

To verify whether the LMDNN model is suitable for making predictions based on many categorical input features, we compared its predictive performance on the IoTP dataset against other conventional machine learning models: recursive partitioning tree [42], conditional inference tree (CTree) [14], random forests and bagging ensemble based on CTree (conditional random forests), SVM, nave Bayes, and logistic regression. We chose these models because they are widely used in privacy decision support systems based on machine learning (e.g., decision trees [12, 36], random forests [35], and SVM [6]). For each machine learning model, we measured its predictive performance on the dataset, through 10-fold cross validation (CV). We only utilized the five contextual parameters as input features in these experiments, without cross-product feature transformations for memorization. This is because we first needed to assess the generalization capability of these models since the dataset is small (6,618 rows), thereby potentially leading to overfitting. Therefore, we evaluated a deep model of LMDNN (see the left side of Fig. 1) in this experiment[1]. Regarding a performance metric, we primarily used the area under the ROC curve (AUC) because it is unaffected by the class imbalance problem (there were 64% allow and 36% reject decisions in the dataset) and is independent of the threshold applied to compute the probability of the binary classification results. Additionally, AUC itself is comprehensible; a random classifier has an AUC score of 0.5 while a perfect classifier has an AUC score of 1.0.

Experimental results indicated that LMDNN outperforms all other models (see Table 1). However, the performance difference is not large (up to 3%). Conditional random forests yield competitive performance because it has proven effective at generalizing to variants not seen in the training set [4], just as deep neural networks do. As mentioned before, however, LMDNN is known to further enhance the deep model by efficiently memorizing feature-target patterns that are rarely observed in a sparse dataset (wide learning). For this distinctive feature, we decided to utilize LMDNN as the machine learning model for this study.

[1] The TensorFlow implementation of LMDNN provides an API that enables programmers to selectively configure a wide, deep, or wide and deep model.

Machine Learning Model	AUC
Recursive Partitioning Tree	0.6142
Conditional Inference Tree	0.6293
Conditional Random Forests	0.6353
Support Vector Machine	0.6107
Naïve Bayes	0.6161
Logistic Regression	0.6208
Deep Neural Networks (of LMDNN)	**0.6421**

Table 1: Predictive Performance of ML Models (10-fold CV)

4.2 Features

Here we explained how we identified the most useful features for training LMDNN models. Specifically, we presented the reasons why we chose the five contextual parameters with interactions between them as underlying input features. In addition, we described how we conducted privacy segmentation on the study participants in [24] and why we utilized the resulting privacy segment information as an additional feature.

4.2.1 Contextual Information

We decided to use the five contextual parameters as basic features for the following reasons: (1) our previous research [24] indicated that all these contextual parameters impact people's privacy decision-making, and (2) we aimed to make predictions based on contextual information which can be automatically collected by IoT environments, thereby avoiding as much as possible asking users to manually enter additional information. We also considered interactions between the contextual parameters because people's privacy decisions about specific contextual information could be influenced by other factors. For instance, the monitoring of personal information (e.g., face photos: C_2=11) can be perceived differently depending on who is performing this monitoring (e.g., unknown vs. employer/school: C_3=1 vs. 6). Furthermore, LMDNN models can make predictions for unusual feature-target patterns by considering cross-product feature transformations (i.e., memorization). Because the contextual parameters *what* (C_2) and *who* (C_3) have the most significant influence on people's privacy decisions [23, 24], we determined that we should feed feature transformations based on *what* and *who* parameters (e.g., $C_2 \times C_1$, $C_2 \times C_3$, $C_2 \times C_4$, $C_2 \times C_5$) into the LMDNN models.

4.2.2 Privacy Segment

As discussed before, privacy segmentation is known to provide useful information for understanding and predicting people's privacy behavior. By applying the K-modes clustering algorithm on the IoTP dataset, we identified distinctive privacy segments and then assigned each user into one of the segments. Even though K-means clustering is the most famous data mining technique, we could not directly apply K-means to the categorical IoTP dataset because it can only process continuous numerical values as its input. As a variant of K-means, the K-modes clustering algorithm was designed to cluster categorical (or ordinal) values without data conversions. It modifies the original K-means by (1) replacing cluster means with cluster modes, (2) using the simple matching dissimilarity function instead of Euclidean distance to calculate the

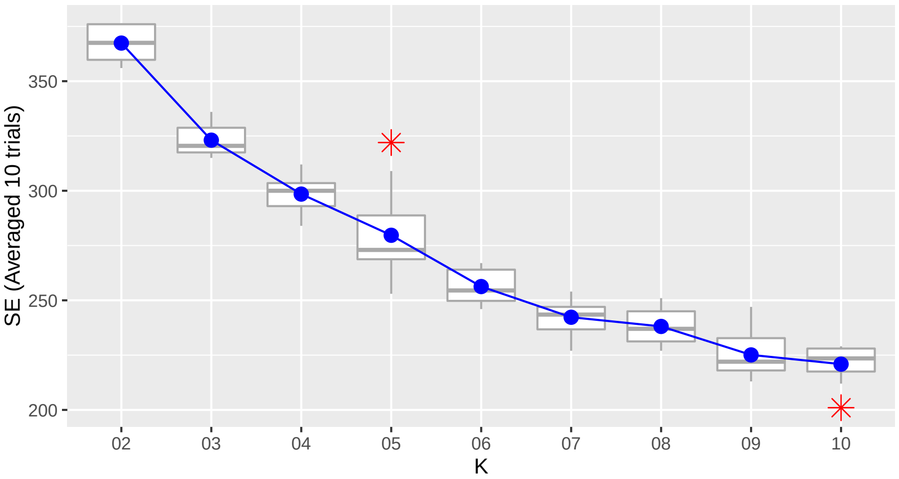

Figure 2: Clustering Errors (*SE*) and the Size of Clusters (*K*)

distance between categorical objects, and (3) updating modes with the most frequent categorical values in each iteration of the clustering [15, 16]. We used klaR [34], an R implementation of K-modes, for the task of privacy segmentation. In [24], we had already performed cluster analysis on the same dataset. However, in that study, we clustered 130 IoT *scenarios* in terms of users' privacy concerns about contextual factors (e.g., *what* and *who*). In this study, we aimed to cluster *users* into privacy segments based on each user's expectation of privacy in a single IoT scenario.

To that end, we selected scenario #60[2] as a base scenario that all participants had responded to, and filtered the dataset by considering the base scenario only. All participants responded to scenario #60 because it is related to the location where each experiment started [24]. By analyzing this partial dataset, we expected to understand how individual users perceive and react differently in the same scenario. To determine the optimal number of clusters (K) for this new dataset, we used the well-known Elbow method [29]. First, we computed the sum of errors (*SE*) of the K-modes clustering with a maximum of 50 iterations, while increasing K from 2 to 10. The *SE* is defined as the sum of the distance between each instance of the cluster and the cluster's centroid (mode). We repeated this procedure 10 times, and took average values of *SE* for each value of K. Next, we calculated the values for the mean difference between SE_K and $SE_{K\text{-}1}$, and found that the largest decrease in errors occurred when we increased K from 2 to 3 (see Fig. 2). Therefore, we chose 3 as a suitable number of clusters, and used it as a parameter (modes) for running the K-modes clustering algorithm on the dataset.

Table 2 summarizes the resulting cluster modes, which are composed of both contextual and reaction parameter values. Note that all contextual parameter values are identical across the clusters because we fixed them to describe the base scenario.

[2] A device of ICS (C_3=6) takes a photo of you (C_2=11). This happens once (C_5=0), while you are in DBH (C_1=3), for safety purposes (C_4=1), namely to determine if you are a wanted criminal.

Contextual Param $\{C_1,C_2,C_3,C_4,C_5\}$	Reaction Param $\{R_1,R_2,R_3,R_4,R_5\}$	Privacy Segment (Cluster Mode)
3, 11, 6, 1, 0	1, **1**, 4, 6, 6	Indifferent (M_1)
3, 11, 6, 1, 0	1, **0**, 3, 3, 3	Some Sensitive (M_2)
3, 11, 6, 1, 0	1, **0**, 1, 1, 1	Sensitive (M_3)

Table 2: Modes of the Resulting Privacy Segments (Clusters)

The clusters are quite distinct from each other, primarily in the reaction parameters *_comfort* (R_3), *_risk* (R_4, reverse-coded), and *_appropriateness* (R_5): each mode has a unique combination of values for R_3, R_4, and R_5. As shown in Table A2, these parameters represent people's privacy attitudes about IoT scenarios on a scale of 1 to 7. For example, cluster mode 3 (M_3) contains R_3=1, R_4=1, and R_5=1, indicating that the given scenario is perceived by participants as *very uncomfortable*, *very risky*, and *very inappropriate*, respectively. For M_3, the value of the reaction parameter *_permission* (R_2) is zero. This means that users belonging to cluster 3 are likely to reject the base scenario because they have negative views on privacy in this scenario. Therefore, we marked the clusters (privacy segments) using these three reaction parameters. We labeled privacy segment 1 (PS_1) as *Indifferent to Privacy* since its mode contains the second highest value for R_4 and R_5 (namely, 6 on a 7-item scale). Likewise, we labeled privacy segment 2 (PS_2) as *Somewhat Sensitive to Privacy* (the values of R_3, R_4, and R_5 fall slightly below the scale average), and privacy segment 3 (PS_3) as *Sensitive to Privacy*. As a result, 74%, 15%, and 11% of the participants were assigned into *Indifferent to Privacy*, *Somewhat Sensitive to Privacy*, and *Sensitive to Privacy* segments, respectively. Finally, we repeated this clustering on additional scenarios (#20, #73, #93, #111)[3] which were the next most frequently visited scenarios after #60. We arrived at the same conclusions regarding the number (K=3) and labels of the resulting privacy segments.

To validate the distinctiveness of the resulting privacy segments, we performed two Welch's t-tests on the R_3 parameter between the following pairs of privacy segments: (PS_1, PS_2) and (PS_2, PS_3). The reason for using Welch's t-test is that all privacy segments have different variances in the R_3 parameter. The tests confirm that the difference in the means of the R_3 parameter between each pair of the segments is statistically significant ($p < 0.025$, Bonferroni-corrected for two comparisons). Next, we also conducted Welch's t-tests on the R_4 and R_5 parameters and drew the same conclusion. Thereby, we verified that the privacy segments are sufficiently distinct from each other in terms of participants' reactions to the given scenario.

Because privacy segment information for all users was available from such a cluster analysis, we then utilized it as an additional feature for building predictive models. This is because privacy segment information is known to be useful for quantifying an individual's judgement about privacy and utility in various circumstances [20, 22].

4.2.3 Predictive Performance Evaluation

As described in Section 4.1.2, we treated the five contextual parameters as basic features for training the deep learning part of LMDNN. We then considered interrelated parameters, especially based on the *what* (C_2) and *who* (C_3) parameters, in conducting both wide and deep learning via LMDNN. This is not only because both contex-

[3] Number of respondents (scenario ID): 140 (#20), 138 (#73), 136 (#93), 162 (#111)

Feature Combination	AUC
(1) C_1, C_2, C_3, C_4, C_5	0.6421
(2) $C_1, C_2, C_3, C_4, C_5, C_2 \times C_{\{1,3,4,5\}}, C_3 \times C_{\{1,4,5\}}$	0.6528
(3) $C_1, C_2, C_3, C_4, C_5, C_2 \times C_{\{1,3,4,5\}}, C_3 \times C_{\{1,4,5\}}, PS$	**0.6725**

Table 3: Predictive Performance of LMDNN Models (10-fold CV)

tual parameters have a significant impact on people's privacy decisions, but because the memorization of some interactions between parameters (i.e., cross-product feature transformations) can improve the performance of the deep model of LMDNN. As an additional feature, we adopted privacy segment information because it differentiates users according to their perceptions of privacy. We tested the following combinations of features to assess their influence on the predictive performance of LMDNN models.

1. Contextual parameters (*deep* learning)
2. Contextual parameters with interactions (*wide* and *deep* learning)
3. Contextual parameters with interactions and privacy segment information (*wide* and *deep* learning)

Using the whole IoTP dataset, we performed 10-fold CV on the LMDNN model trained with each of these feature combinations. As expected, the AUC score gradually improves as we added cross-product feature transformations and privacy segment information to the five basic contextual features (see Table 3). We then concluded that both the contextual parameters (including interactions) and privacy segment information can act as informative features for predicting the binary value of the reaction parameter *_permission* (R_2) through LMDNN.

4.3 Training Strategy

Based on the selected machine learning model (LMDNN) and features (contextual and privacy segment information), we investigated the best strategy to build a predictive model (i.e., classifier) for each individual user. First, we reviewed two commonly used strategies, individual and one-size-fits-all modeling. Individual modeling typically utilizes a single user's instances as training data, and will therefore result in a highly personalized user-specific model if a sufficient amount of training data is available. One-size-fits-all modeling, in contrast, trains a single model based on all users' data, so that reasonable predictions can be made about new users for whom insufficient data is available. Next, we proposed our strategy that we dub one-size-fits-segment modeling. It was designed to utilize the one-size-fits-all paradigm for making predictions for users grouped by privacy segmentation. We analyzed the overall and per-user performance of the predictive models trained through these three different strategies, and then draw some practical implications.

4.3.1 Individual Modeling

This is the most popular and straightforward approach for building user-specific machine learning models for privacy decision support systems, especially targeted at predicting users' decisions about the disclosure of personal information on social network services [35, 12, 36, 37, 40, 6]. It constructs a distinctive predictive model per each user by using his/her data only. This assumes that each user has a very different point of view regarding privacy, therefore the others' data are not useful for modeling and predicting his/her own privacy decision-making. As most previous works have adopted supervised machine learning approaches, each user will need to provide a certain amount of labeled training data (e.g., historical decisions for privacy-invasive scenarios). For instance, Bilogrevic et al. [6] stated that they needed 40 manual decisions from each user to build personalized models with a reasonable predictive performance, which is a quite burdensome amount. The more training data a user provides, the more the performance of the individual predictive model tends to improve. For these reasons, an individual modeling strategy is suitable for situations in which service providers could acquire enough training data from each individual user. Like [6], the study participants in [24] made about 40 privacy decisions (reaction parameter *_permission*) on average. Therefore, we first applied individual modeling for constructing personalized LMDNN models to check whether this strategy is adequate for solving our own problem. Specifically, we trained and evaluated a single LMDNN model for a specific user based on his/her data. We repeated it for all users in the dataset. As input features, we utilized solely the five contextual parameters with their interactions. Since each individual LMDNN model is exclusively trained using each user's data, we do not use privacy segment information in this experimental setup.

4.3.2 One-size-fits-all Modeling

One-size-fits-all contrasts with an individual modeling strategy. Instead of building multiple user-specific predictive models, it generates a single universal model based on data collected from a crowd of users. Because a one-size-fits-all predictive model is trained using a larger dataset (multiple users' data), it typically represents a wider range of common feature-target patterns than an individual model. Therefore, researchers have often utilized a one-size-fits-all modeling strategy for building a predictive model that makes initial predictions (e.g., default privacy settings) for new users who did not provide training data [6]. Recently published works [41, 30] also adopted one-size-fits-all modeling to make their privacy decision support systems generalizable to a wide range of users. The authors in [30] collected IoT-related privacy decisions from 1,007 Amazon Turkers through an online survey, and built one-size-fits-all predictive models for randomly chosen 50 participants. They reported that the overall accuracy of these trained models ranges from 76% to 80%, depending on whether most (75%) or all (100%) of the other participants' responses are used as training data. To verify the applicability of the one-size-fits-all modeling strategy to privacy decision support for real-world IoT environments (using IoTP dataset), we built a one-size-fits-all LMDNN model for each user based on data collected from all other users, utilizing both contextual and privacy segment information as input features. Each user's data was used thereafter to assess the predictive power of his/her LMDNN model.

4.3.3 One-size-fits-segment Modeling

We proposed a novel approach called one-size-fits-segment modeling. This is a modified version of one-size-fits-all modeling. It builds multiple universal models for several groups of users rather than for the entire user population. Here, we utilized the result of privacy segmentation (see Section 4.2.2) as a criterion to cluster users. We intended to improve the performance of one-size-fits-all modeling by constructing per-user predictive models based on data collected from same-minded users in terms of privacy. We believe that each individual user will be benefitted if a predictive model is

Training Strategy	Mean AUC	Std. Dev.
Individual	0.4806	0.2179
One-size-fits-all	0.6699	0.1721
One-size-fits-segment	**0.6782**	**0.1668**

Table 4: Predictive Performance of LMDNN Models

trained on large volumes of data provided by others similar to him/her. To verify the proposed approach, we conducted experiments as follows. First, we divided users in the IoTP dataset into three groups according to our privacy segmentation. For each single user, we determined the privacy segment to which he/she belongs, and built a one-size-fits-segment LMDNN model by utilizing data collected from other users in the corresponding privacy segment. Each user's data was used as test data for measuring the predictive performance of his/her LMDNN model. Unlike one-size-fits-all modeling, we only utilized contextual parameters (including interactions between them) as input features. As explained, privacy segment information was utilized to split users with regard to the subset of their stated privacy decisions, thereby constructing one-size-fits-segment LMDNN models for every single user. In comparison with individual modeling and one-size-fits-all modeling (described in [24]), the proposed mechanism does not burden new users because it just asks five questions (reaction parameters) about a single base scenario to perform privacy segmentation.

4.3.4 Predictive Performance Evaluation

We compared the predictive power of these three different training strategies as follows. Regarding individual modeling, we trained and assessed the predictive performance of user-specific LMDNN models via 10-fold CV. Each of these models was trained on all data instances from a different user in the IoTP dataset. We repeated this procedure for all users and took the average of their AUC scores as the overall predictive performance of individual modeling. For one-size-fits-all or one-size-fits-segment modeling, we constructed a target user's LMDNN model based on all other users' data and tested the trained model using the target user's data. Unlike [30], we repeated this for all users and calculated the average AUC score for each strategy. Finally, we compared these mean AUC scores for three training strategies to choose the best.

Table 4 summarizes the experimental results showing that both the one-size-fits-all and the one-size-fits-segment modeling strategy significantly outperform the individual modeling strategy. Furthermore, one-size-fits-segment shows a slightly better performance than one-size-fits-all. This is because one-size-fits-segment LMDNN models are trained by the data of same-minded users, thereby predicting each user's privacy decisions more accurately. This finding was unexpected since most previous research about privacy decision support has reported that individual modeling is better than one-size-fits-all modeling for inferring users' privacy decision-making. One possible explanation is that the size of per-user training data was not large enough to learn sparse high-dimensional feature spaces. Participants in the dataset responded to an average of 39.58 scenarios (std. dev: 14.65), or about 33% of all available scenarios. Each scenario received 50.9 responses on average (std. dev: 40.16). When we transformed the dataset into a user-scenario matrix (i.e., each cell represents an observed feature-target pattern), the sparsity[4] of this matrix was 0.6566. This low response rate may also be

[4] 1 - (nonzero entries/total entries in a user-scenario matrix)

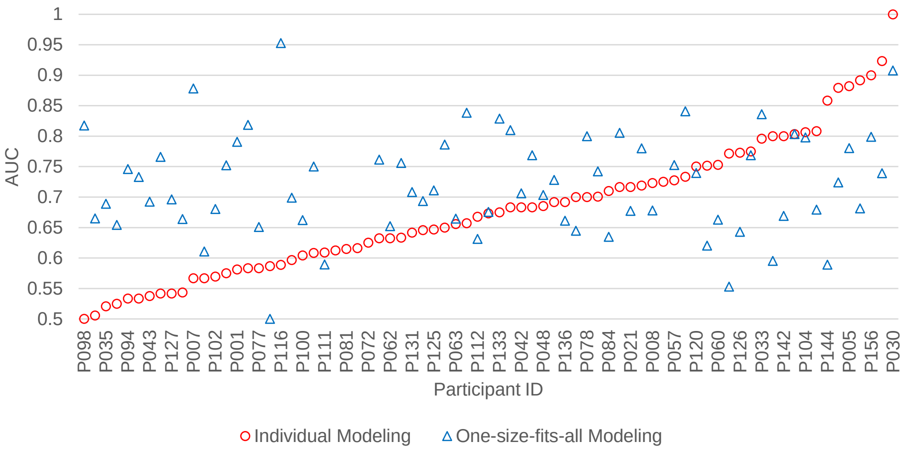

Figure 3: Per-user Predictive Performance

found in future real-world situations since the number of applications and services in IoT environments is likely to increase over time. As a result, it could be difficult to collect enough training data from each IoT user to build individual predictive models from scratch.

Fig. 3 presents the per-user predictive performance of two model training strategies: individual and one-size-fits-segment modeling. Results have been filtered out by the threshold of $AUC > 0.5$ and sorted by the AUC score of individual LMDNN models in ascending order. As can be seen, one-size-fits-segment LMDNN models show a better predictive performance than individual LMDNN models for about 80% of the users (left side of the chart). However, there is also the opposite effect for about 20% of users (far right side of the chart) who have highly accurate individual LMDNN models ($AUC > 0.7$) trained solely on their data. To check the causes of this per-user difference, we performed Pearson's chi-square tests for independence between personal attributes (gender, age, type of university affiliation)[5] and the fitted training strategy. For instance, we constructed a 2×2 contingency table using two binary variables, age (18-25 or older than 25) and training strategy (one-size-fits-segment or individual modeling), and assessed the significance of the difference between the two proportions (i.e., users aged 18-25 and older users, both have superior one-size-fits-segment LMDNN models). The test confirms that there is no statistical evidence of an association between these two variables at the .05 significance level. We iterated this test for other personal attributes such as gender (male or others) and the type of university affiliation (undergraduate students or others), and reached the same conclusion. Identifying latent factors (e.g., cultural background in privacy decision-making [26]) that cause this difference is one possible future research direction. If we were to find these factors, we could determine the most appropriate modeling strategy for each individual user in advance, and then accordingly utilize it for better predictive performance.

[5] Mode values of these attributes are male, 18-25, and undergraduate students, respectively.

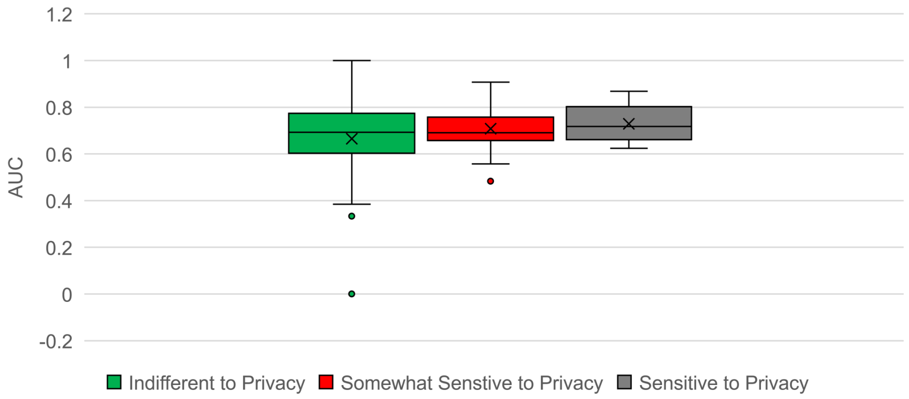

Figure 4: Per-segment Predictive Performance

4.4 Implications

We confirmed that privacy segment information is helpful for understanding and predicting people's privacy decision-making about various IoT services. Specifically, one-size-fits-segment modeling shows the best predictive performance (mean AUC of 0.6782; N=172) compared with the previously proposed model training strategies. However, even a single *false positive* may cause undesired information monitoring in IoT, making users reluctant to use privacy decision support systems. At its current accuracy, we should therefore use the prediction results only to give recommendations to users that they can still inspect and override; we should not use the prediction results for automated disclosure decisions. Moreover, we should strive for further improvement of the predictive performance. In the following, we explained how the current predictive performance could be improved.

4.4.1 Role of Privacy Segment Information

To better interpret the performance of one-size-fits-segment modeling, we measured the predictive performance of the trained LMDNN models for each privacy segment. Fig. 4 indicates that users who are assigned to the *Sensitive to Privacy* segment have the most accurate LMDNN models (mean AUC of 0.729, std. dev: 0.074). Likewise, users belonging to the *Somewhat Sensitive to Privacy* segment also tend to have LMDNN models with a reasonable performance (mean AUC of 0.708, std. dev: 0.101). Users in the *Indifferent to Privacy* segment however have relatively inaccurate LMDNN models compared to those in other privacy segments (mean AUC of 0.665, std. dev: 0.184). Here, we also could not identify any inter-segment differences in terms of personal characteristics such as gender, age, and the type of university affiliation. One possible explanation is that participants in both the *Sensitive to Privacy* and *Somewhat Sensitive to Privacy* segments presumably responded to IoT monitoring scenarios more carefully than participants in the *Indifferent to Privacy* segment, thereby providing more consistent privacy decisions (i.e., high-quality training data). We might need to consider strategies for improving the data collection process for the *Indifferent to Privacy* segment. For instance, users who are indifferent about privacy could be asked for more training data than the other users. This is because a higher amount of training data would reduce uncertainty in predictive models, and therefore likely improve performance. It might also be possible to utilize privacy nudges [28], which can make privacy-insensitive users aware of unexpected outcomes of personal information monitoring, for steering users towards more consistent privacy behavior.

4.4.2 Size of Individual Training Data

Regarding individual modeling, 28 users had highly accurate ($AUC > 0.7$) individual LMDNN models trained on their own data. Furthermore, individual models can be continuously improved as each user provides additional training data. To verify this claim, we built a linear regression model with the number of user responses (i.e., size of training data) as the independent variable and AUC scores (i.e., predictive performance) as the dependent variable. According to the fitted regression model, a statistically significant positive relationship exists between these two variables ($p < 0.005$). Therefore, the collection of extra training data might further improve the performance of individual machine learning models. Yet, asking for additional data can also be burdensome and privacy-invasive.

4.4.3 Hybrid Modeling

For these reasons, it might be desirable to gradually morph one-size-fits-segment models into individual models. Like [6, 41, 30], we verified that our one-size-fits-segment LMDNN models can make reasonably accurate predictions for two segments, even for new users. Therefore, we can utilize them to produce *default* privacy settings that are applicable to the general population (grouped by privacy segmentation). As mentioned before, however, these general predictive models would need to be tailored to each user to make more accurate predictions. One possible avenue would be to gradually transform one-size-fits-segment LMDNN models into *personalized* individual LMDNN models for each user. Specifically, it is necessary to continuously retrain (update) the base one-size-fits-segment LMDNN model with user-provided training data while increasing the weights of this user-provided data (i.e., transfer learning [32, 18, 25]). This should then better fit the updated LMDNN model with the user's unique behavioral patterns. Active learning paradigms can also be applied to reduce users' labeling efforts as much as possible.

5 Discussion and Limitations

We studied machine learning mechanisms for modeling and predicting users' privacy decisions regarding personal information tracking in IoT environments. After investigating various machine learning algorithms, input features, and model training strategies, we proposed a novel mechanism called one-size-fits-segment modeling for privacy decision support systems in IoT. We showed that the proposed mechanism exhibits a reasonable predictive performance on privacy behavioral data captured in the field. Our work goes far beyond the limited machine learning approach in [24], where contextual parameters (without consideration for interaction effects) and cluster membership information of IoT service scenarios were used as features for building a decision tree-based predictive model. We believe this approach is not very practical because it requires users to specify reaction parameter values for dozens of scenarios in advance. Another difference between [24] and this work lies in the examination of the model training strategies; we only tested one-size-fits-all modeling (i.e., single classifier) in [24], and hence the results may provide limited implications (e.g., lack of consideration for per-user predictive performance). Yet, our work still has some shortcomings that will be addressed in the next few sub-sections.

5.1 Representability of Data

First, we need to consider the representativeness of the dataset we used. The study participants were predominantly university students aged 18-25 (82%), since we recruited them on campus. This may induce a sampling bias that makes our results less generalizable. In other words, we do not know how the proposed mechanism will work on datasets collected from other populations (e.g., older users who are not familiar with IoT). In this regard, we plan to validate our mechanism with more representative samples, thereby confirming a future direction of this research (e.g., continuous updates of one-size-fits-segment LMDNN models with user-provided data). It will also be necessary to apply the proposed mechanism to different domains (e.g., privacy decisions in healthcare settings), and verify its expandability. To that end, we need to collect or get access to users' privacy behavioral data regarding such a domain.

5.2 Reliability of Privacy Segmentation

We utilized the results of privacy segmentation to improve the performance of our predictive models. As discussed, we performed privacy segmentation by clustering users into several groups based on their responses toward a single scenario. Here, we utilized scenario #60 as a base scenario since all users responded to this scenario; this enabled us to assign privacy segment information to each of the users. We also tried four different base scenarios to validate whether our clustering methodology (i.e., determining the correct number and labels of the clusters) is sound, and the result was positive. However, it is also important to check the invariance of the clustering results. As future work, we plan to collect all users' privacy decisions regarding more than one base scenario and conduct pair-wise comparisons on the clustering results so as to confirm the resulting privacy segments are stable no matter which base scenario is used for privacy segmentation. If a user was assigned to the same privacy segment regardless of the base scenario, we can consider the results of privacy segmentation as ground truth. Otherwise, we need to devise a way to decide on the most accurate privacy segment for this user. One possible approach is a majority vote among the segmentation results from all base scenarios.

5.3 Privacy Paradox

The privacy paradox is the phenomenon that people's stated privacy preferences or decisions often seem inconsistent with their actual behaviors [2, 17, 11]. As explained before, we utilized stated privacy decisions collected through ESM in a simulated IoT environment [24] as training data, and not actual behavior observed in a working IoT environment. Although we had tried to make participants perceive they were in a real situation, we do not know how they would actually behave in real-world situations. Therefore, we need to construct a working IoT environment, let participants freely interact with the environment, and then collect the corresponding privacy decisions, possibly with sensor data. To that end, we plan to conduct an experimental study for collecting (large-scale) privacy behavioral data from real users in an operational

IoT system based on open protocols for IoT discovery and interaction, such as the Open Connectivity Foundation (OCF), Web of Things, or Physical Web. By using this dataset, we need to confirm the effectiveness of the proposed mechanism.

6 Conclusion

In this paper, we proposed a novel machine learning mechanism for predicting people's privacy decisions in IoT environments. We aimed to predict binary privacy decisions for each user, namely whether to allow or reject a given personal information monitoring scenario in IoT. To begin with, we adopted linear model and deep neural networks (LMDNN) as the machine learning model for our study. Using a privacy behavioral dataset (N=172) collected from our earlier study, we confirmed that LMDNN provides better predictive performance than conventional machine learning models that have been widely used in the literature. Next, we utilized both contextual and privacy segment information as input features for training LMDNN models. We adopted a wide range of contextual factors comprising diverse IoT scenarios. We then generated users' privacy segment information by clustering their privacy decisions about a single selected IoT scenario. Lastly, we proposed a new model training strategy called one-size-fits-segment modeling, and compared its performance with two commonly used strategies: individual and one-size-fits-all modeling. Experimental results indicated that one-size-fits-segment outperforms other modeling strategies. We also presented some practical implications regarding the design and development of privacy decision support systems for IoT environments. Future work will focus on collecting privacy-related human behavioral data from more representative samples of users interacting with working IoT systems, and validating the proposed mechanism on this new dataset.

Acknowledgement

This research was funded by NSF grant SES-1423629. The human subjects research described herein is covered under IRB protocol #2014-1600 at the University of California, Irvine.

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.

[2] Alessandro Acquisti and Jens Grossklags. Privacy attitudes and privacy behavior. In *Economics of Information Security*, pages 165–178. Springer, 2004.

[3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805, 2010.

[4] Yoshua Bengio, Olivier Delalleau, and Clarence Simard. Decision trees do not generalize to new variations. *Computational Intelligence*, 26(4):449–467, 2010.

[5] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing*, 15(7):679–694, 2011.

[6] Igor Bilogrevic, Kévin Huguenin, Berker Agir, Murtuza Jadliwala, Maria Gazaki, and Jean-Pierre Hubaux. A machine-learning based approach to privacy-aware information-sharing in mobile social networks. *Pervasive and Mobile Computing*, 25:125–142, 2016.

[7] Grégoire Burel, Hassan Saif, and Harith Alani. Semantic wide and deep learning for detecting crisis-information categories on social media. In *International Semantic Web Conference*, pages 138–155. Springer, 2017.

[8] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pages 7–10. ACM, 2016.

[9] Richard Chow, Serge Egelman, Raghudeep Kannavara, Hosub Lee, Suyash Misra, and Edward Wang. HCI in business: A collaboration with academia in IoT privacy. In *International Conference on HCI in Business*, pages 679–687. Springer, 2015.

[10] Delphine Christin, Andreas Reinhardt, Salil S Kanhere, and Matthias Hollick. A survey on privacy in mobile participatory sensing applications. *Journal of Systems and Software*, 84(11):1928–1946, 2011.

[11] Kay Connelly, Ashraf Khalil, and Yong Liu. Do I do what I say?: Observed versus stated privacy preferences. *Human-Computer Interaction–INTERACT 2007*, pages 620–623, 2007.

[12] Lujun Fang and Kristen LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th international conference on World Wide Web*, pages 351–360. ACM, 2010.

[13] Christine Hine. Privacy in the marketplace. *The Information Society*, 14(4):253–262, 1998.

[14] Torsten Hothorn, Kurt Hornik, and Achim Zeileis. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3):651–674, 2006.

[15] Zhexue Huang. A fast clustering algorithm to cluster very large categorical data sets in data mining. *DMKD*, 3(8):34–39, 1997.

[16] Zhexue Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, 2(3):283–304, 1998.

[17] Carlos Jensen, Colin Potts, and Christian Jensen. Privacy practices of Internet users: self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1):203–227, 2005.

[18] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing image style. *arXiv preprint arXiv:1311.3715*, 2013.

[19] Fatemeh Khatibloo. It's here! forrester's consumer privacy segmentation. Available at `https://go.forrester.com/blogs/its-here-forresters-consumer-privacy-segmentation/` (2016/12/15).

[20] Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. Dimensionality of information disclosure behavior. *International Journal of Human-Computer Studies*, 71(12):1144–1162, 2013.

[21] Ponnurangam Kumaraguru and Lorrie Faith Cranor. Privacy indexes: a survey of Westin's studies. Technical report, Carnegie Mellon University, Pittsburgh, PA, 2005.

[22] Nancy Lankton, D McKnight, and John Tripp. Privacy management strategies: An exploratory cluster analysis. In *Proceedings of the 22nd Americas Conference on Information Systems (AMCIS 2016)*, pages 1–10, 2016.

[23] Hosub Lee and Alfred Kobsa. Understanding user privacy in Internet of Things environments. In *Internet of Things (WF-IoT), 2016 IEEE 3rd World Forum on*, pages 407–412. IEEE, 2016.

[24] Hosub Lee and Alfred Kobsa. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on*, pages 276–285. IEEE, 2017.

[25] Hosub Lee, Cameron Upright, Steven Eliuk, and Alfred Kobsa. Personalized object recognition for augmenting human memory. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 1054–1061. ACM, 2016.

[26] Yao Li, Alfred Kobsa, Bart P Knijnenburg, Carolyn Nguyen, et al. Cross-cultural privacy prediction. *Proceedings on Privacy Enhancing Technologies*, 2017(2):113–132, 2017.

[27] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I Hong. Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *Proceedings of the 10th Symposium on Usable Privacy and Security (SOUPS 2014)*, pages 199–212, 2014.

[28] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, SA Zhang, Norman Sadeh, Alessandro Acquisti, and Yuvraj Agarwal. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Proceedings of the 12th Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 27–41, 2016.

[29] T Soni Madhulatha. An overview on clustering methods. *arXiv preprint arXiv:1205.1117*, 2012.

[30] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. Privacy expectations and preferences in an IoT world. In *Proceedings of the 13th Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 399–412, 2017.

[31] Patricia A Norberg, Daniel R Horne, and David A Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007.

[32] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[33] Charith Perera, Rajiv Ranjan, Lizhe Wang, Samee U Khan, and Albert Y Zomaya. Big data privacy in the Internet of Things era. *IT Professional*, 17(3):32–39, 2015.

[34] Christian Roever, Nils Raabe, Karsten Luebke, Uwe Ligges, Gero Szepannek, and Marc Zentgraf. Package klar. Available at `https://cran.r-project.org/web/packages/klaR/klaR.pdf/` (2015/2/20).

[35] Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, 2009.

[36] Mohamed Shehab, Gorrell Cheek, Hakim Touati, Anna C Squicciarini, and Pau-Chen Cheng. User centric policy management in online social networks. In *Policies for Distributed Systems and Networks (POLICY), 2010 IEEE International Symposium on*, pages 9–13. IEEE, 2010.

[37] Mohamed Shehab and Hakim Touati. Semi-supervised policy recommendation for online social networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*, pages 360–367. IEEE, 2012.

[38] Shaoyun Shi, Min Zhang, Hongyu Lu, Yiqun Liu, and Shaopin Ma. Wide & deep learning in job recommendation: An empirical study. In *Asia Information Retrieval Symposium*, pages 112–124. Springer, 2017.

[39] Sabrina Sicari, Alessandra Rizzardi, Luigi Alfredo Grieco, and Alberto Coen-Porisini. Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, 76:146–164, 2015.

[40] Arunesh Sinha, Yan Li, and Lujo Bauer. What you want is not what you get: Predicting sharing policies for text-based content on Facebook. In *Proceedings of the 2013 ACM Workshop on Artificial Intelligence and Security*, pages 13–24. ACM, 2013.

[41] Eleftherios Spyromitros-Xioufis, Georgios Petkos, Symeon Papadopoulos, Rob Heyman, and Yiannis Kompatsiaris. Perceived versus actual predictability of personal information in social networks. In *International Conference on Internet Science*, pages 133–147. Springer, 2016.

[42] Terry M Therneau, Elizabeth J Atkinson, et al. An introduction to recursive partitioning using the RPART routines. Technical report, Mayo Foundation, 1997.

A Appendix

Parameter (id)	Values
where (*C_1*)	(0) your place (1) someone else's place (2) semi-public space (e.g., restaurant) (3) public space (e.g., street)
what (*C_2*)	(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→lookingAt (23) gaze (24) gaze→lookingAt
who (*C_3*)	(1) unknown (2) colleague/fellow (3) friend (4) own device (5) business (6) employer/school (7) government
reason (*C_4*)	(1) safety; (2) commercial; (3) social; (4) convenience; (5) health; (6) none
persistence (*C_5*)	(0) once (1) continuously

Table A1: Contextual Parameters

Parameter (id)	Values
_notification (R_1)	(1) notify me, always
	(2) notify me, just this time
	(3) don’t notify me, just this time
	(4) don’t notify me, always
_permission (R_2)	(1) allow, always
	(2) allow, just this time
	(3) reject, just this time
	(4) reject, always
_comfort (R_3)	(1) very uncomfortable
	(2) uncomfortable
	(3) somewhat uncomfortable
	(4) neutral
	(5) somewhat comfortable
	(6) comfortable
	(7) very comfortable
_risk (R_4)	(1) very risky
	(2) risky
	(3) somewhat risky
	(4) neutral
	(5) somewhat safe
	(6) safe
	(7) very safe
_appropriateness (R_5)	(1) very inappropriate
	(2) inappropriate
	(3) somewhat inappropriate
	(4) neutral
	(5) somewhat appropriate
	(6) appropriate
	(7) very appropriate

Table A2: Reaction Parameters