Given social media users’ plethora of interactions, appropriately controlling access to such information becomes a challenging task for users. Selecting the appropriate audience, even from within their own friend network, can be fraught with difficulties. PACMAN is a potential solution for this problem. It’s a personal assistant agent that recommends personalized access control decisions based on the social context of any information disclosure by incorporating communities generated from the user’s network structure and utilizing information in the user’s profile. PACMAN provides accurate recommendations while minimizing intrusiveness.

Social media has become synonymous with communication in daily life for most of us. On Facebook alone, more than 1 billion users share over 300 petabytes of personal information daily. Social media users interact with people representing various facets of their life, such as work, family, and education. In such a scenario, it’s essential for them to make informed access control decisions to preserve the “contextual integrity” of their information. Any user who discloses information on social media has a notion of the “intended recipients” and the context in which they would view that information, and hence, preservation of “contextual integrity” is essential to avoid a privacy breach. Unfortunately, the privacy controls afforded to users by social media sites make it burdensome to selectively share content within their friend network; this results in a situation where they end up sharing their information with “unintended recipients.”

The mainstream social media sites such as Facebook and Google+ have taken steps to mitigate this by assisting users in managing their friend networks by creating Lists and Circles, respectively. However, recent research findings suggest that hardly any users employ these features when making access control decisions, arguably because of the effort this requires from them.

Social media users can be assisted by recommendation systems that can guide them toward the appropriate access control decisions. It’s well-established that different users exhibit different access control behavior and often have differing privacy preferences. Therefore, it’s essential that a
recommendation system forms the core of a personal agent that can provide personalized recommendations to individual users. In recent times, we’ve seen personal agents being proposed to provide assistance to users in various social media issues, such as ascertaining contexts of disclosure, detecting privacy violations when they happen, and negotiating multiparty privacy conflicts. However, to the best of our knowledge, there’s an absence of a personal agent that recommends personalized access control decisions to users to minimize the burden of expressing their individual sharing preferences.

Thus, here we present PACMAN, a personal agent that provides a novel approach to learning access control decisions by combining social relationships and information about the content. The building blocks of PACMAN are identified by conducting detailed empirical evaluations that result in a highly accurate mechanism using a minimal set of appropriate attributes. Our results show that PACMAN produces an average accuracy of 91.8 percent (standard deviation = 6.5 percent, median = 94.1 percent) across all users. We find that PACMAN works best for users who are more static in terms of the number of friends to whom they grant access.

**PACMAN**

Figure 1 shows the information that PACMAN uses to produce an access control recommendation (“allow” or “deny”). For social media users, the social context of information disclosure is considered essential to enable the formulation of access control policies in a way that preserves the information’s “contextual integrity.” The social context can be derived from information that facilitates the definition of social relationships on these media. These interpersonal relationships can be defined in terms of relationship types, often denoted by communities and the relationship strength or closeness, which is represented by the similarity of profile attributes. In addition to relationships, the content itself is an integral part of the context of the disclosure and plays an important role in the formulation of a desired access control policy.

Therefore, the information about the content being shared (text, photos, and so on) also should be used to learn or determine access control decisions.

**Relationship Type**

Social media users have various types of interpersonal relationships (such as friends, colleagues, and family) with the people they interact with on the network. These can often be represented by partitioning one’s network into groups or communities. These partitions can be leveraged by any access control mechanism such that the user might be asked to make access control decisions with respect to one or some members in a particular “community” (created by the algorithm), and then implement that decision for the other members in that community. PACMAN uses network-based community detection and requires the user’s friend network as an input.

**References**

Relationship Strength

The interpersonal relationships between social media users can also be defined in terms of strength (or closeness). This is generally estimated by measuring similarity between individuals’ profiles. There have been several proposed approaches in literature that suggest appropriate methods of estimating tie-strength or closeness, such that it can be used to assist users in making informed access control decisions. However, they all have the same limitations: First, the information required from the profiles might not be easy to fetch and process, which makes it difficult to provide users with real-time assistance on a dynamic medium such as social media; second, some profile attributes often are missing, as users often refrain from populating many fields on their social media profiles; and third, accessing certain types of personal information from the users’ profiles might be intrusive, and hence, counterproductive for a privacy-preserving mechanism. In our previous work, we performed a systematic analysis of all profile attributes available in social media profiles to select the minimal subset most suitable for predicting access control decisions with maximum possible accuracy. The analysis led to the identification of Total Friends (the total size of a user’s friend network) and Mutual Friends (the number of shared friends or contacts with the user) as the most appropriate profile attributes to enable prediction of access control decisions, while overcoming the discussed challenges. Therefore, PACMAN uses these two attributes to account for the relationship strength between a user and each of his or her friends.

Content

The information about the content being shared also can be used to enhance access control mechanisms. The information about the content can be automatically mined and used to classify the content, which then can be leveraged to inform access control decisions. Different methods can be used to create attributes, depending on the content’s nature (for example, natural language processing techniques can be used for text, and image processing can be used for photos). However, such analysis is still far from being completely automated in terms of accuracy to represent the user’s perception about the content. One method of mitigating this is by asking users to provide metadata in the form of “tags” while sharing the content. Previous research shows that such tags also can be used to create access control policies and that they’re minimally disruptive for the user. PACMAN is agnostic to the type of content being shared, and hence, different methods of obtaining information about the content can be implemented. If automatic analysis of content is implemented in PACMAN, it can operate completely without any user input, because PACMAN automatically analyzes the other attributes, representing relationship type and relationship strength.

Evaluations

To evaluate whether and to what extent access control decisions made by social media users can be learned by PACMAN, we conducted a user study to obtain ground truth access control decisions to use for learning, which is the standard way of evaluating automated access control mechanisms in the literature.

Experiment

We created an application using Facebook Query Language (FQL) and the Facebook Graph API for participants to make access control decisions while disclosing 10 photos. Five of these photos were randomly downloaded from their Facebook profiles, and the participants were asked to select and bring five other photos that they hadn’t uploaded to Facebook yet, to avoid a scenario where a user makes access control decisions for all photos during the study for which they had already received comments and likes before, as that might have influenced their decisions. The
participants were advised to bring photos that they considered to be personal (which either included them or a family member) or considered sensitive so that they had a privacy implication. The different stages were as follows:

- The participants logged into the application using their Facebook credentials. They were then alerted about the data that would be accessed and asked for explicit permissions before moving on.
- The participants were shown 10 photos sequentially on the screen, each on an individual page. They were asked to select categories for the photos from a predefined list of 15 popular Flickr categories, and make access control decisions for each photo. The friend list was shown alphabetically to the participants and they were instructed to select each and every friend that they would want to grant access to the photo. They were explicitly informed that any friend who wasn’t selected would be denied access to the photo.
- Once participants made the access control decisions and selected the categories for all 10 photos, their selections, friend lists, and Total Friends and Mutual Friends profile attributes of all their friends were stored.

**Participants**

This research experiment was conducted at Lancaster University after being approved by the university’s Research Ethics Committee. Participants were recruited primarily from among the university’s staff and students. Additionally, we invited some participants who were external to the university through personal communication channels such as email and social networks. Each participant was compensated £10 for being in the study.

We applied the typical pre- and post-experiment checks to maximize data quality. In particular, before the experiment we screened participants and everyone who had a Facebook account and had uploaded at least 10 photos before the study was eligible to participate. After an initial registration phase, 31 participants took part in the study. After completion of the user study, we checked all responses to make sure participants had correctly completed the experiment, finding five participants who didn’t (four had randomly selected lists of alphabetically sorted friends, and one had selected one single but different friend for each photo). The remaining 26 participants were considered for analyses, including 15 males (57.7 percent) and 11 females (42.3 percent). The average participant’s age was 29 years (standard deviation = 6) and the average social network size was 265 friends (standard deviation = 121). The total number of access control decisions made by the 26 participants during the experiment, and hence, the size of the ground truth dataset, was 67,660.

**Implementing PACMAN**

As we described earlier, we implemented PACMAN’s design using various building blocks to represent the different components shown in Figure 1. The information required from users’ Facebook profiles and their access control decisions were obtained from the user study as described.

To represent the Relationship Type, PACMAN uses community membership. In our previous work,17 we evaluated eight well-known network-based community detection algorithms for a goodness of fit with access control decisions made by social media users. Our analysis found Clique Percolation Method (CPM) to be the most suitable community detection algorithm in an access control scenario and CPM membership is used to represent the Relationship Type in this implementation of PACMAN. The friend network of each user obtained during the user study was used as input to the CPM algorithm, which was implemented using the iGraph library to create communities. Each of a user’s friends was assigned a community membership that was denoted using a binary vector, with dimension equal to the total communities of the user, to represent their relationship type in PACMAN. For this implementation, we used non-overlapping CPM communities such that each of the users’ friends belonged to exactly one community.

For Relationship Strength, the Total Friends and Mutual Friends attributes were directly fetched from the users’ profiles during the study and used as input to the PACMAN mechanism.

As mentioned earlier, PACMAN’s design is agnostic to the type of content being shared as well as the method used to obtain information about the content. In this particular implementation, we used manual selection of photo categories in the form of “tags” to represent the
information about content. This was done as it provided us with the user’s perspective about the content in a comparatively less-intrusive way.\textsuperscript{16} The users during the study were given an opportunity to select categories for the photos in the form of tags, as mentioned earlier. While it wasn’t mandatory to select categories for each photo, we found that only 4 out of the 260 photos (10 per user) weren’t categorized. The average number of categories selected per photo was found to be 2.2. The content information was represented with a binary vector having a dimension of 15 (the total number of categories) representing whether each category was selected. Thus, a photo which wasn’t categorized would be represented as all zeroes.

For evaluating PACMAN’s performance, \textit{Weka} was integrated into PACMAN to create and run the classifier using 10-fold cross validation to calculate accuracy of prediction produced for each individual user. There were 67,660 instances in total, corresponding to all the access control decisions in the ground truth dataset. The attributes consisted of the CPM membership vector, total, and mutual friends, as well as the content vector representing the photo categories. In 10-fold cross-validation, the entire dataset is randomly divided into 10 subsets, each of which are then used as training data (while leaving the other nine as test sets) for each iteration. This process is repeated 10 times such that each subset gets to be the training set and the average error across all 10 iterations is considered as the final value. We performed 10-fold cross-validation using the in-built function present in \textit{Weka}, which automatically divides the dataset into 10 random subsets. To the best of our knowledge, this is the most rigorous and systematic method of evaluating a classifier, because it rules out the possible bias associated with division of a dataset into training and test sets.

PACMAN can work with any machine learning algorithm and for the evaluation, we tried a \textit{Naive Bayes} classification algorithm, support vector machines (SVM), and Random Forest, but found that Random Forest produced the best results and have only reported those in this article because of space constraints.

\textbf{Estimating User Effort}

PACMAN recommends “allow” or “deny” access control decisions to the user, corresponding to each member in their friend network. For PACMAN, both classes “allow” and “deny” are of equal importance, as users would spend time and effort in correcting the erroneous recommendations made by PACMAN. In such a scenario, accuracy is appropriate, as other metrics focus on giving more importance to one of the classes;\textsuperscript{18} for example, when a program is to be classified as malware or not, positive classification is prioritized. To calculate accuracy, we take the access control decisions made by users for all 10 photos during the user study as the ground truth. In particular, for a user having $F$ total friends, PACMAN’s accuracy can be calculated as a percentage of the total recommendations that are correct:

\begin{equation}
\text{Accuracy} = \frac{(F - \text{Errors})}{F} \tag{1}
\end{equation}

The \textit{Errors} include both “allow” and “deny” errors.

An \textit{Allow error} occurs when PACMAN recommends a “deny” decision to the user when it actually should have been “allow.” These errors are essentially “false negative” (FN) recommendations and result in a “deny to allow” change being made by the user.

A \textit{Deny error} occurs when PACMAN recommends an “allow” decision to the user when it actually should have been “deny.” These errors are “false positive” (FP) recommendations and result in an “allow to deny” change by the user.

\begin{equation}
\text{Errors} = \text{FN} + \text{FP} \tag{2}
\end{equation}

We show the ratio of both types of errors for each user to provide a more precise picture of PACMAN’s performance, regarding each type of error.

In addition to reporting the accuracy of the recommendations made by PACMAN, we also show the area under ROC curve (AUC; ROC stands for receiver operating characteristic) to give an idea of the quality of recommendations made by PACMAN.

\textbf{Results}

Now that we’ve discussed an implementation of PACMAN, we describe the results of our analyses in this section.

\textbf{Overall Accuracy}

Figure 2 shows the accuracy of recommendations produced by PACMAN for each of the 26 users. It also shows the ratio of incorrect
recommendations, *Allow errors* and *Deny errors*. We can see in Figure 2 that PACMAN produces highly accurate recommendations for almost all users. The average accuracy across 26 users was found to be 91.8 percent (standard deviation = 6.5 percent, min = 81.4 percent, max = 99.7 percent). We find that almost all users have similar amounts of *Allow errors* (mean = 5.4 percent, standard deviation = 3.9 percent) and *Deny errors* (mean = 2.8 percent, standard deviation = 3.5 percent). This suggests that PACMAN doesn’t discriminate between the two classes and that recommendation errors are fairly equal. The AUC was 0.845 (standard deviation = 0.097), which shows PACMAN produces good-quality recommendations.

### Clustering Users

To enhance our understanding of PACMAN’s strengths and weaknesses, we wanted to examine the factors which might distinguish users for whom it produces high accuracy, as compared to the ones with comparatively lower accuracy. We used two-step clustering, using overall accuracy as the clustering variable, to obtain the two clusters of users as described in Table 1.

We find a cluster of 17 users for whom PACMAN produces very high accuracy (mean = 96.1 percent, standard deviation = 2.9 percent). These users have a comparatively more static access control behavior, with a lower average and standard deviation for audience sizes (across 10 photos), and a smaller number of communities. The nine users in the other cluster were found to have comparatively lower — but still decent — accuracy (mean = 83.8 percent, standard deviation = 1.9 percent). It’s noticeable that they have greater variation in their access control behavior, with higher average and standard deviation for the audience sizes and number of communities. Table 1 also shows that both clusters have similarly high AUC values, which suggests that PACMAN produces good-quality recommendations for all users.

We also calculated the correlation coefficients with respect to accuracy and the access control behavior of users. These coefficients are shown in Table 2. The correlations confirm the hypothesis that users who have larger average audiences and larger variations in their selections are more likely to have higher errors (both *Allow errors* and *Deny errors*) and a comparatively lower accuracy as a result.

We didn’t find any significant correlations in terms of the personal characteristics of the users, such as gender, age, number of photos uploaded (amount of activity on Facebook), or size of the friend network. No significant trends could be observed with respect to the category or source (Facebook or USB) of photos in terms of accuracy of PACMAN prediction. This suggests that PACMAN would work for all...
categories of photos and whether they had been uploaded previously on social media doesn’t have an effect on its performance.

**Contribution of Types of Attributes**

We wanted to examine whether all three types of attributes were required and were contributing to the performance of PACMAN, or whether one or more were redundant and could be avoided without compromising performance. We calculated the relative information gain for each type of attribute as a ratio of the total information gain to compare the contribution for each individual user.

Table 1 shows the aggregated values for all 26 users as well as both clusters of users. The numbers suggest that all components contribute to the performance of PACMAN while relationship strength seems to contribute the most for the average user. The difference between the clusters wasn’t found to be statistically significant. Nevertheless, the numbers suggest that PACMAN relies more on the content for users who select larger audiences and have greater variation. Therefore, it’s plausible that the PACMAN accuracy for such users would improve by training with more photos for each type of photo content.

Our personal assistant agent, PACMAN, leverages information about interpersonal relationships between individuals on social networks and combines this with information...
about the content to recommend access control decisions. Our evaluations show that PACMAN produces highly accurate access control recommendations, and all three components of PACMAN are important — each individual component has varying importance for different users. Interestingly, PACMAN tends to rely more on content for users who select larger audiences and have greater variation in their access control behavior.

Having considered only network-based community detection for representing relationship types in PACMAN, we can consider social circles — based on contextual information beyond social media profiles, such as co-location — as a possible future enhancement. We can use sensors on mobile devices to identify contacts in the same location, and then use this information as an attribute. Looking at the reliance of PACMAN on content for users with greater variation in access control behavior, other methods of extracting information about content such as the physical properties of the photos themselves could be considered to observe whether it enhances the accuracy for such users. This would enable PACMAN to function without any user input and make it work in a scenario where a social network is a network of agents that make access control decisions based on automatic analysis of the attributes. Finally, PACMAN focuses on learning individual preferences, which also could be used as input to other tools that recommend access control decisions for multiuser scenarios.

References

Gaurav Misra recently completed his PhD research at Lancaster University, UK. His doctoral research focused on creating solutions to access control problems faced by social media users. He is now working as a postdoctoral research fellow at the University of New South Wales (UNSW) in Canberra, Australia, where he tackles problems emanating from the social aspects of security and privacy. Contact him at g.misra@adfa.edu.au.

Jose M. Such is a senior lecturer (associate professor) in computer science at King’s College London. His research interests are at the intersection between cybersecurity, artificial intelligence, and human-computer interaction, with a strong focus on privacy, intelligent access control, and co-owned data in socio-technical and cyber-physical systems. Such has a PhD in computer science from the Universitat Politecnica de Valencia. Contact him at jose.such@kcl.ac.uk.