PACMAN: Personal Agent for Access Control in Social Media

考虑到用户过多的交互，恰当地控制对这些信息的访问成为用户的具有挑战性的任务。 选择适当的受众，即使是在他们自己的朋友网络中，也可能充满了困难。 PACMAN是这个问题的潜在解决方案。 它是一个私人助理代理，根据任何信息披露的社会背景，通过合并从用户的网络结构生成的社区，并利用用户信息中的信息，推荐个性化的访问控制决策。 PACMAN提供准确的建议，同时最大限度地减少侵入性。

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.

Gaurav Misra
Lancaster University

Jose M. Such
King's College London

社交媒体已经成为我们大多数人日常生活中沟通的代名词。仅在Facebook上，每天就有超过10亿用户共享超过300 PB的个人信息。用户与代表他们生活各个方面的人们（如工作，家庭和教育）进行交互。在这种情况下，他们必须做出明智的访问控制决策，以保持其信息的“上下文完整性”。任何披露关于a的信息的用户都有“预期接收者”的概念以及他们将查看该信息的上下文，因此保留“上下文完整性”对于避免隐私泄露是必不可少的.1不幸的是，隐私控制由网站提供给用户使得选择性地在他们的朋友网络内共享内容成为麻烦;这导致他们最终以“无意的收件人”分享他们的信息。2 Facebook和Google1等网站的主流已经采取措施，通过分别创建“列表”和“圈子”来帮助用户管理好友网络。然而，最近的研究结果表明，几乎没有任何用户在进行访问控制决策时使用这些功能，这可能是因为这需要他们的努力。

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.[1] 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."[2] 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.[3]

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

 Published by the IEEE Computer Society 1089-7801/17/$33.00 © 2017 IEEE

Related Work in Determining Access Control Decisions

There have been many previous works in the area of predicting and recommending access control decisions to social media users. Many of these works use different types of features to enable prediction of access control decisions. Some approaches advocate the user of community membership, while others rely on profile information[2,3] to recommend access control decisions. The information about the content is also used to determine the appropriate audience to share it with.[4] PACMAN advances the state of art by using a combination of relationship-based attributes, communities, and profile attributes to represent the "who," and the information about the content (the "what") to represent the social context of the disclosure.

References

1. L. Fang and K. LeFevre, "Privacy Wizards for Social Networking Sites," *Proc. 19th Int'l Conf. World Wide Web*, 2010, pp. 351–360.
2. S. Amershi, J. Fogarty, and D. Weld, "Regroup: Interactive Machine Learning for On-Demand Group Creation in Social Networks," *Proc. Sigchi*, 2012, pp. 21–30.
3. C. Akcora, B. Carminati, and E. Ferrari, "Privacy in Social Networks: How Risky Is Your Social Graph?" *Proc. 28th Int'l Conf. Data Eng.*, 2012, pp. 9–19.
4. A.C. Squicciarini et al., "A3P: Adaptive Policy Prediction for Shared Images over Popular Content Sharing Sites," *Proc. 22nd ACM Conf. Hypertext and Hypermedia*, 2011, pp. 261–270.

推荐系统可以帮助社交媒体用户，这些推荐系统可以指导他们走向适当的访问控制决策。 不同的用户表现出不同的访问控制行为，并且通常具有不同的隐私偏好。因此，推荐系统必须构成个人代理的核心，可以为个人用户提供个性化的推荐。近来，我们已经看到私人代理商在各种问题上提供帮助，比如确定披露的背景，检测隐私侵犯事件发生的时间，以及协商多方隐私冲突.6但是， 在我们的知识中，缺少一个向用户推荐个性化访问控制决策的个人代理，以最小化表达他们个人共享偏好的负担。

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,[4] detecting privacy violations when they happen,[5] and negotiating multiparty privacy conflicts.[6] 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.

因此，在这里，我们介绍PACMAN，一个私人代理，通过结合社会关系和内容信息，提供一种学习访问控制决策的新方法。 PACMAN的构建块通过进行详细的经验评估来确定，这些评估导致使用最少的适当属性的高度精确的机制。我们的结果显示，PACMAN的平均准确率为91.8%（标准偏差5 6.5%，中位数5 94.1%）。 我们发现PACMAN最适合那些在授权访问的朋友数量方面更加静态的用户。

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."[7] 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[8] and the *relationship strength* or closeness, which is represented by the similarity of profile attributes.[9,10] 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.

图1显示了PACMAN用来产生访问控制建议（"允许"或"拒绝"）的信息。 对于社交媒体用户而言，信息披露的社会背景被认为是必不可少的，以保持信息的"语境完整性"的方式来制定访问控制政策。7社会背景可以从促进社会关系定义的信息 这些媒体。 这些人际关系可以根据关系类型来定义，通常由社区来表示[8]，关系的强度或亲密度由表达特征的相似度来表示[9,10]。除了关系外，内容本身也是关系的一个组成部分 并且在制定期望的访问控制策略中起重要作用。

因此，也应该使用关于共享内容的信息（文本，照片等）来学习或确定访问控制决策。

Therefore, the information about the content being shared (text, photos, and so on) also should be used to learn or determine access control decisions.[11]

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 then 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.[12] PACMAN uses network-based community detection and requires the user's friend network as an input.

社交媒体用户与他们在网络上互动的人有各种人际关系（如朋友，同事和家人）。 这些通常可以通过将一个人的网络划分成组或社区来表示。 这些分区然后可以被任何访问控制机制所利用，使得可以要求用户针对特定"社区"（由算法创建）中的一个或一些成员做出访问控制决定，然后实施针对该社区中的其他成员.12 PACMAN使用基于网络的社区检测，并要求用户的朋友网络作为输入。

关于被共享的内容的信息也可以用来增强访问控制机制。有关内容的信息可以自动挖掘并用于对内容进行分类，然后利用这些信息来通知访问控制决策.15根据内容的性质，可以使用不同的方法来创建属性（例如，自然语言处理技术可以用于文字，图像处理可以用于照片）。然而，就精确性而言，这样的分析还远未完全自动化以代表用户对内容的感知。减轻这一点的一种方法是要求用户在共享内容的同时以“ 标签 ”的形式提供元数据。以前的研究表明，这样的标签也可以用来创建访问控制策略，并且对用户的干扰最小。[16] PACMAN对共享内容的类型是不知道的，因此，获取内容信息的不同方法可以执行。如果在PACMAN中实现内容的自动分析，它可以在没有任何用户输入的情况下完全运行，因为PACMAN自动分析其他属性，表示关系类型和关系强度。

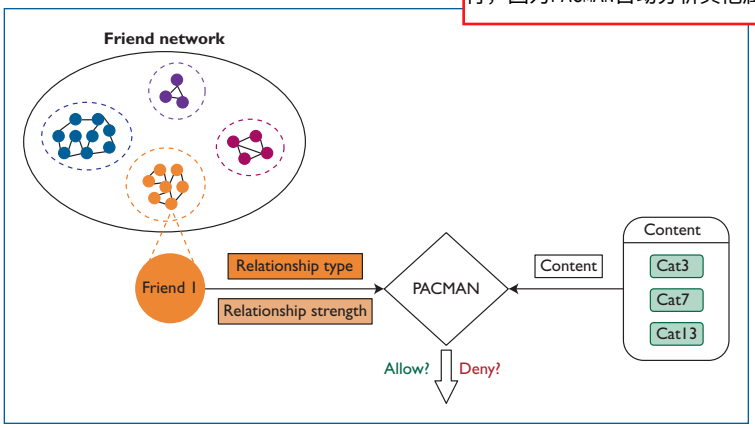


Figure 1. Components and inputs of PACMAN, an agent for social media that offers personalized access control.

Relationship Strength

社交媒体用户之间的人际关系也可以用强度（或亲密度）来定义。通常通过测量个人的特征之间的相似性来估计这一点。文献中提出了几种建议方法，提出了适当的估计联系强度或亲密度的方法，这样可以帮助用户做出明智的访问控制决策.13但是，它们都具有相同的局限性：首先，信息所需要的信息可能不易获取和处理，这使得用户在诸如社交媒体等动态媒体上提供实时帮助变得困难。其次，一些特征属性经常缺失，因为用户通常不会在他们的社交媒体上使用许多长子;第三，从用户的特征中获取特定类型的个人信息可能是侵入性的，因此对于隐私保护机制是适得其反的。在我们以前的工作中，我们对社交媒体专业中可用的所有特征属性进行了系统分析，以最大可能的精度选择最适合预测访问控制决策的最小子集。分析导致了对全部朋友（用户的朋友网络的总大小）和相互朋友（与朋友分享的朋友的数量或与用户的联系）的识别作为最合适的特征属性以使得能够预测访问控制决策，同时克服讨论的挑战。因此，PACMAN使用这两个属性来说明用户与他或她的每个朋友之间的关系强度。

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.[13] 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;[9] 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,[14] 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.[15] 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.[16] 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.

为了评估社交媒体用户的访问控制决策能否被PACMAN学习到，以及在何种程度上可以被PACMAN学习到，我们进行了一项用户研究，以获得用于学习的地面真实访问控制决策，这是评估自动访问控制机制的标准方法 文献。

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

我们使用Facebook查询语言（FQL）和Facebook图形API创建了一个应用程序，供参与者进行访问控制决策，同时披露10张照片。 其中五张照片是从他们Facebook的脸书上随机下载的，参与者被要求选择并带来五张尚未上传到Facebook的照片，以避免用户在 他们之前已经收到过评论和喜欢的研究，因为这可能会影响他们的决定。 建议参与者带上他们认为是个人的照片（包括他们或家庭成员）或被认为是敏感的，以便他们有隐私的含义。 不同的阶段如下：

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. ifferent stages were as follows:

•参与者使用他们的Facebook凭证登录到应用程序。 然后，他们被告知将要访问的数据，并在继续之前被要求获得明确的许可。
•参与者在屏幕上顺序显示了10张照片，每张照片都在一个单独的页面上。 他们被要求从15个流行的Flickr类别的预定义列表中选择照片类别，并为每张照片做出访问控制决定。 朋友列表按字母顺序显示给参与者，他们被指示选择他们想要授予访问照片的每个朋友。 他们被明确告知，任何未被选中的朋友将被拒绝访问照片。
•一旦参与者做出访问控制决策并为所有10张照片选择类别，则存储他们的所有朋友的选择，朋友列表以及所有朋友的共同朋友和共同朋友特征属性。

e participants logged into the application ing their Facebook credentials. They were en alerted about the data that would be cessed and asked for explicit permissions fore moving on.

e participants were shown 10 photos quentially on the screen, each on an indi- dual page. They were asked to select cate- ries for the photos from a predefined list of popular Flickr categories, and make access ntrol decisions for each photo. The friend t was shown alphabetically to the partici- nts and they were instructed to select each d every friend that they would want to ant access to the photo. They were explic- y informed that any friend who wasn't ected would be denied access to the photo. ce participants made the access control cisions and selected the categories for all photos, their selections, friend lists, and tal Friends and Mutual Friends profile ributes 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.

这项研究实验是经过大学研究伦理委员会批准后在兰开斯特大学进行的。 参与者主要来自大学的工作人员和学生。 此外，还通过电子邮件和社交网络等个人交流渠道，邀请了一些外部参与者到大学。每个参与者在研究中获得了10英镑的补偿

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.

我们应用了典型的实验前和实验后检查来最大化数据质量。特别是在实验之前，我们筛选了参与者和每个拥有Facebook账号的人，并在研究前至少上传了10张照片。在初始登记阶段，31名参与者参加了这项研究。在完成用户研究之后，我们检查了所有回答，以确保参与者已经正确地完成了实验，找到了五个没有参与的人（四个随机选择的按字母排序的朋友列表，一个为每张照片选择了一个但不同的朋友）。其中26人被认为是男性，其中男性15人（57.7%），女性11人（42.3%）。平均参与者为29岁（标准偏差5 6），平均社交网络规模为265朋友（标准差5 121）。实验中26名参与者进行的访问控制决策的总数量，以及地面真实数据集的大小为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.

正如我们前面所描述的那样，我们使用各种构建块来实现帕克曼的设计，以表示图1中所示的不同组成部分。用户的脸谱和他们的访问控制决策所需的信息是从所描述的用户研究中获得的。

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.

为了表示关系类型，PACMAN使用社区成员。 在我们以前的工作中，我们评估了八种基于网络的知名社区检测算法，以及一个由社交媒体用户做出的访问控制决策。 我们的分析发现，在访问控制场景和CPM成员中，最适合的社群检测算法是派生渗透法（CPM），用于表示PACMAN实现中的关系类型。用户研究中获得的每个用户的好友网络被用作输入 到使用iGraph库实现社区的CPM算法。 每个用户的朋友被分配了一个社区成员资格，用二进制向量表示，维度等于用户的总社区，以表示他们在PACMAN中的关系类型。 对于这个实现，我们使用了非重叠的CPM社区，比如用户的朋友属于一个社区

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.

对于"关系强度"，"全友好友"和"共同朋友"的属性直接从学习期间的用户特征中提取出来，并用作PACMAN机制的输入。

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

PACMAN建议对用户的“允许”或“拒绝”访问控制决定，对应于他们的朋友网络中的每个成员。 对于PACMAN，“允许”和“否认”两个类别都是重要的，因为用户会花时间去纠正PACMAN提出的错误建议。 在这种情况下，准确性是适当的，因为其他度量标准更注重对其中一个类的重视；例如，当一个程序被分类为恶意软件时，正分类被优先考虑。 为了计算准确度，我们将用户在学习期间所有10张照片的用户访问控制决策作为基础事实。 特别是，对于有F个朋友的用户，PACMAN的准确度可以计算为正确的总推荐的百分比：

如前所述，PACMAN的设计对共享内容的类型以及用于获取内容信息的方法是不可知的。在这个特定的实现中，我们用“标签”的形式手动选择光子类别来表示关于内容的信息。这样做是因为它提供了用户对相对较少干扰内容的内容的观点.16如前所述，研究期间的用户有机会以标签的形式为照片选择类别。虽然这不是强制性的为每张照片选择类别，我们发现在260张照片中只有4张（每个用户10张）没有被分类。每张照片选择的平均类别数量是2.2。内容信息被表示为具有代表每个类别是否被选择的15维（类别的总数）的二元向量。因此，未被分类的照片将被表示为全零。

nformation about content. This was done as it provided us with the user's perspective about the content in a comparatively less-intrusive way.[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.

为了评估PACMAN的性能，Weka被整合到PACMAN中，使用10倍交叉验证来创建和运行分类器，以计算为每个用户生成的预测的准确性。共有67,660个物质，对应于地面实况数据集中的所有访问控制决策。属性包括CPM成员向量，总数，共同朋友以及代表照片类别的内容向量。在10倍交叉验证，整个数据集被随机分成10个子集，每个子集作为训练数据（而另外9个作为测试集）用于每次迭代。重复这个过程10次，使每个子集成为训练集，并且所有10次迭代的平均误差被认为是最终值。我们使用Weka中的内置函数进行10倍交叉验证，自动将数据集划分为10个随机子集。据我们所知，这是评估分类器的最严格和系统的方法，因为它排除了与将数据集划分为训练集和测试集相关的可能偏差。PACMAN可以和任何机器学习算法一起使用，并且对于评估，我们尝试了一种朴素贝叶斯分类算法，支持向量机（SVM）和随机森林，但是发现随机森林产生了最好的结果，并且只报告了由于空间约束

For evaluating PACMAN's performance, 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 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 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.

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;[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:

$$Accuracy = ((F - Errors)/F). \quad (1)$$

The *Errors* include both “allow” and “deny” errors.

An *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 *Deny error* occurs when PACMAN re mends an “allow” decision to the user wh actually should have been “deny.” These e are “false positive” (FP) recommendations result in an “allow to deny” change by the

$$Errors = FN + FP.$$

We show the ratio of both types of erro each user to provide a more precise pictu PACMAN's performance, regarding each of error.

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

错误包括“允许”和“拒绝”错误。当PACMAN向用户推荐一个“拒绝”决定，当其实际上应该是“允许”时，允许错误发生。这些错误基本上是“错误否定”（FN）建议和结果在“否认允许”的情况下，由用户进行修改。当PACMAN推荐对用户进行“允许”决定时，拒绝错误发生在实际上应该被“拒绝”时。这些错误是“误报”（FP）建议和结果“允许拒绝”用户的更改。错误5 FN 1 FP。

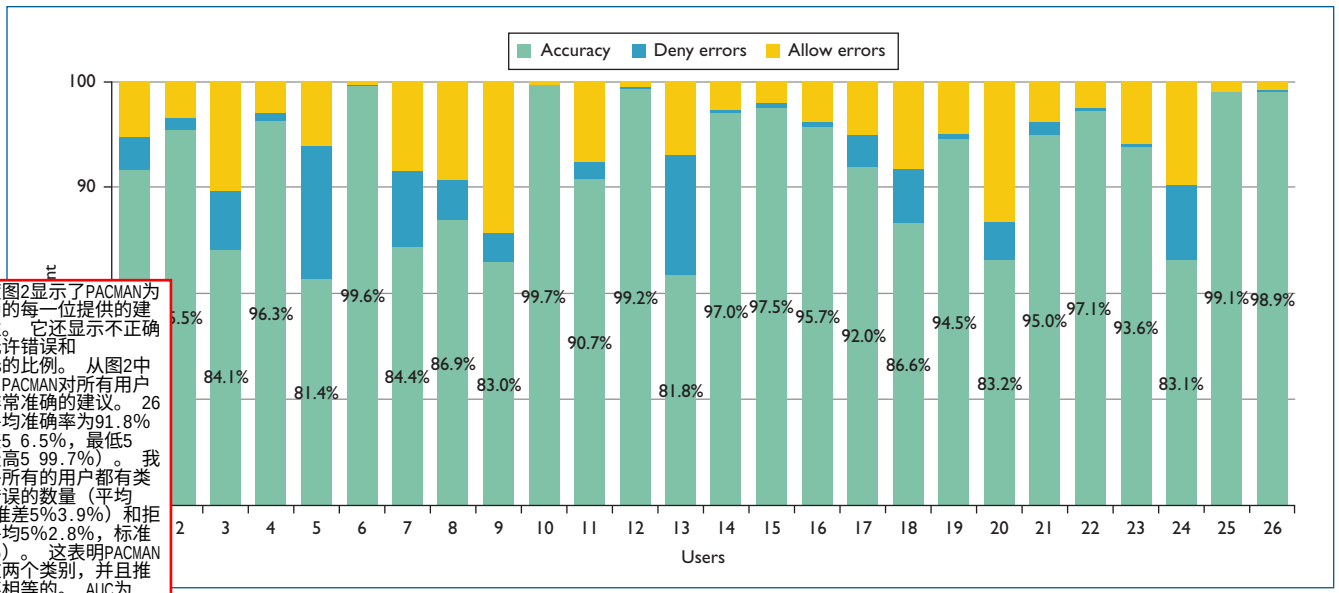
（2）我们给出了两种错误类型之间的比率，以提供关于每种误差的更精确的帕克曼表现的图像。除了报告PACMAN提出的建议的准确性之外，我们还给出了ROC曲线下面积（AUC; ROCstands为接收器操作特性）原谅了PACMAN建议的质量的想法。结果现在我们已经讨论了PACMAN的实现，我们在本节中描述我们分析的结果

Results

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

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



acy and ratio of changes required to recommendations made by PACMAN for all 26 users.

整体准确度图2显示了PACMAN为26位用户中的每一位提供的建议的准确性。它还显示不正确的建议，允许错误和Denyerrors的比例。 从图2中可以看出，PACMAN对所有用户都提供了非常准确的建议。 26个用户的平均准确率为91.8%（标准偏差5 6.5%，最低5 81.4%，最高5 99.7%）。 我们发现几乎所有的用户都有类似的允许错误的数量（平均5.4%，标准差5%3.9%）和拒绝错误（平均5%2.8%，标准差5%3.5%）。 这表明PACMAN不会区分这两个类别，并且推荐错误是不相等的。 AUC为0.845（标准偏差5 0.097），这表明PACMAN提供了良好的质量建议。

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

我们还计算了关于用户的准确性和访问控制行为的相关系数。 表2显示了这些系数。相关性证明了这样的假设，即具有较大的平均观众和较大的选择变化的用户更可能具有较高的误差（允许误差和误差误差）并且由此导致相对较低的精度。
我们没有发现用户的个人特征，如性别，年龄，上传的照片数量（Facebook上的活动量）或朋友网络的大小等方面的显着相关性。 在PACMAN预测的准确性方面，没有观察到关于照片类别或来源（Facebook或USB）的显着趋势。这表明PACMAN将为所有类别的照片工作，并且他们以前是否已经在社交媒体上上载没有影响 在其表现。

用户聚类为了增强我们对PACMAN强度和弱点的理解，我们想要考察哪些因素可以区分用户，从而获得高精度，而精度相对较低。我们使用了两步聚类，以总体精度作为聚类变量，得到如表1所示的用户群。我们找到17个用户群，其中，帕克曼具有很高的准确性（平均值为596.1%，标准差为5%，2.9%），这些用户具有相对较多的静态访问控制行为，对于观众大小（10张照片）具有较低的平均值和标准偏差，以及较少数量的社区。另一个群体中的9个用户被发现相对较低 - 但仍然像样的准确性（平均5 83.8%，标准差5 1.9%）。值得注意的是，他们的访问控制行为有较大的变化，平均值和标准偏差较高表1还显示了两个群集具有相似的高AUC值，这表明PACMAN为所有用户产生高质量的推荐

Table 1. PACMAN accuracy and access control behavior of users in both clusters.

Cluster	Users	Statistic	Average audience	Standard deviation audience	Communities used	Accuracy (%)	Allow errors (%)	Deny errors (%)	AUC*	Relative information gain**		
										Type	Strength	Content
Higher accuracy	17	Average***	15.06	19.65	5.18	96.1	3.1	0.8	0.848	0.170	0.490	0.339
		Standard deviation	14.30	22.34	3.88	2.9	2.2	0.9	0.118	0.290	0.409	0.383
Lower accuracy	9	Average	59.87	57.46	10.33					0.155	0.336	0.509
		Standard deviation	20.61	17.82	7.60					0.162	0.317	0.317
Overall	26	Average	30.57	32.74	6.96					0.165	0.437	0.398
		Standard deviation	27.19	27.52	5.86					0.249	0.381	0.365

* AUC stands for area under ROC curve. ROC stands for receiver operating charac
** The difference in relative information gain values wasn't found to be statistically s ences were found to be statistically significant at the 99 percent confidence interval using the Mann-Whitney test.
*** Average and standard deviation values show aggregate statistics across all users

属性类型的贡献我们想要考察是否所有属性的三类属性都是必需的，并且对PACMAN的性能有贡献，或者是否有一个或多个属性是冗余的，在不影响性能的情况下可以避免。我们计算了每种类型属性的相对信息增益，的总信息收益来比较每个用户的贡献。表1显示了所有26个用户以及两个用户群的汇总值。数字表明，所有组件对PACMAN的性能有贡献，而关系强度似乎对普通用户贡献最大。团体之间的差异没有统计意义上的显着性。尽管如此，这些数字表明，帕克曼更依赖于用户的访问控制行为有较大变化的内容。表2中的相关系数也支持这个观点，因为我们认为，PACMAN更多地依赖内容来选择更多的观众并且有更大的变化。因此，通过对每种类型的照片内容进行更多照片的培训，PACMAN对这些用户的准确性将会得到改善。

Table 2. Pearson correlation of accuracy and contribution of components with access control behavior.

Criteria	Average audience	Standard deviation audience	Communities used
Allow errors	0.660*	0.576*	0.360
Deny errors	0.896*	0.800*	0.636*
Accuracy	−0.880*	−0.777*	−0.558*
AUC	0.198	0.362	0.130
Relative type gain	−0.149	−0.194	0.036
Relative strength gain	−0.312	−0.402**	0.209
Relative content gain	0.428**	0.553*	0.194

* Correlation is significant at the 99 percent confidence level.
** Correlation is significant at the 95 percent confidence level.

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 show greater variation in their access control behavior. This notion is also supported by the correlation coefficients in Table 2, where we find that PACMAN relies more on 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

我们的私人助理代理PACMAN利用社交网络中个人之间的人际关系信息，并将其与关于内容的信息结合起来，以推荐访问控制决策。我们的评估显示，PACMAN提供高度准确的访问控制建议，PACMAN的所有三个组件都非常重要 - 每个组件对于不同的用户具有不同的重要性。有趣的是，PACMAN倾向于更多地依赖于那些选择大成就的用户的内容，并且在访问控制行为上有更大的变化。

在考虑了PACMAN中表示关系类型的基于网络的社区检测之后，我们可以考虑将社交圈以超越社交媒体特征的情境信息（如共同定位）19作为未来可能的增强。我们可以使用移动设备上的传感器来识别同一地点的联系人，然后将这些信息作为解决方案.19从PACMANon内容的依赖性来看，对于具有较大变化的访问控制行为的用户来说，其他方法可以提取有关内容的信息，如物理属性的照片本身11可以考虑是否提高了这些用户的准确性。这将使得PACMAN在没有任何用户输入的情况下运行，并且在社交网络是基于对属性的自动分析进行访问控制决定的代理的网络的情况下使其工作。最后，PACMAN侧重于学习个人偏好，这些偏好也可以用来作为推荐多用户场景的访问控制决策的其他工具。

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[19] — 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.[19] 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[11] 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.[20]

References

1. H. Nissenbaum, “Privacy as Contextual Integrity,” *Washington Law Rev.*, vol. 79, 2004, p. 119.
2. S. Egelman, A. Oates, and S. Krishnamurthi, “Oops, I Did It Again: Mitigating Repeated Access Control Errors on Facebook,” *Proc. Sigchi Conf. Human Factors in Computing Systems*, 2011, pp. 2295–2304.
3. P. Wisniewski, B. P. Knijnenburg, and H. Richter Lipford, “Profiling Facebook Users Privacy Behaviors,” *Proc. SOUPS 2014 Workshop on Privacy Personas and Segmentation*, 2014; https://cups.cs.cmu.edu/soups/2014/workshops/privacy/s2p1.pdf.
4. N. Criado and J.M. Such, “Implicit Contextual Integrity in Online Social Networks,” *Information Sciences*, vol. 325, 2015, pp. 48–69.
5. O. Kafali, A. Gunay, and P. Yolum, “Protoss: A Run Time Tool for Detecting Privacy Violations in Online Social Networks,” *Proc. Int’l Conf. Advances in Social Networks Analysis and Mining*, 2012, pp. 429–433.
8. L. Fang and K. LeFevre, “Privacy Wizards for Social Networking Sites,” *Proc. 19th Int’l Conf. World Wide Web*, 2010, pp. 351–360.
9. S. Amershi, J. Fogarty, and D. Weld, “Regroup: Interactive Machine Learning for On-Demand Group Creation in Social Networks,” *Proc. Sigchi*, 2012, pp. 21–30.
10. J.J. McAuley and J. Leskovec, “Learning to Discover Social Circles in Ego Networks,” *Advances in Neural Information Processing Systems*, vol. 272, 2012, pp. 548–556.
11. A.C. Squicciarini et al., “Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites,” *IEEE Trans. Knowledge and Data Eng.*, vol. 27, no. 1, 2015, pp. 193–206.
12. G.P. Cheek and M. Shehab, “Human Effects of Enhanced Privacy Management Models,” *IEEE Trans. Dependable and Secure Computing*, vol. 11, no. 2, 2014, pp. 142–154.
13. R.L. Fogues et al., “Bff: A Tool for Eliciting Tie Strength and User Communities in Social Networking Services,” *Information Systems Frontiers*, 2013, pp. 1–13.
14. G. Misra, J.M. Such, and H. Balogun, “Improve—Identifying Minimal PROfile VEctors for Similarity Based Access Control,” *Proc. 2016 IEEE Trustcom/BigDataSE/ISPA*, 2016, pp. 868–875.
15. A.C. Squicciarini et al., “A3p: Adaptive Policy Prediction for Shared Images over Popular Content Sharing Sites,” *Proc. 22nd ACM Conf. Hypertext and Hypermedia*, 2011, pp. 261–270.
16. C.-M.A. Yeung et al., “Providing Access Control to Online Photo Albums Based on Tags and Linked Data,” *Proc. AAAI Spring Symp.: Social Semantic Web: Where Web 2.0 Meets Web 3.0*, 2009, pp. 9–14.
17. G. Misra, J. M. Such, and H. Balogun, “Non-Sharing Communities? An Empirical Study of Community Detection for Access Control Decisions,” *Proc. IEEE/ACM Int’l Conf. Advances in Social Networks Analysis and Mining*, 2016, pp. 49–56.
18. J.L. Herlocker et al., “Evaluating Collaborative Filtering Recommender Systems,” *ACM Trans. Information Systems*, vol. 22, no. 1, 2004, pp. 5–53.
19. P. Murukannaiah and M. Singh, “Platys Social: Relating Shared Places and Private Social Circles,” *IEEE Internet Computing*, vol. 16, no. 3, pp. 53–59, 2012.
20. R.L. Fogues et al., “Sharing Policies in Multiuser Privacy Scenarios: Incorporating Context, Preferences, and Arguments in Decision Making,” *ACM Trans. Computer-Human Interaction*, in press, 2017.

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.