Assisting Users in a World Full of Cameras
A Privacy-aware Infrastructure for Computer Vision Applications

Anupam Das, Martin Degeling, Xiaoyou Wang, Junjue Wang, Norman Sadeh, Mahadev Satyanarayanan
Carnegie Mellon University

Abstract

Computer vision based technologies have seen widespread adoption over the recent years. This use is not limited to the rapid adoption of facial recognition technology but extends to facial expression recognition, scene recognition and more. These developments raise privacy concerns and call for novel solutions to ensure adequate user awareness, and ideally, control over the resulting collection and use of potentially sensitive data. While cameras have become ubiquitous, most of the time users are not even aware of their presence. In this paper we introduce a novel distributed privacy infrastructure for the Internet-of-Things and discuss in particular how it can help enhance user’s awareness of and control over the collection and use of video data about them. The infrastructure, which has undergone early deployment and evaluation on two campuses, supports the automated discovery of IoT resources and the selective notification of users. This includes the presence of computer vision applications that collect data about users. In particular, we describe an implementation of functionality that helps users discover nearby cameras and choose whether or not they want their faces to be denatured in the video streams.

1. Introduction

Computer vision has been an active field of research for many decades and as a result a myriad of applications have been developed using computer vision based technologies [43]. Examples of computer vision based technologies that we encounter in our everyday life are gesture recognition (e.g., Kinect), image search (e.g., Google images), facial recognition (e.g., Facebook’s Moment app) and automated cars or bots (e.g., Uber cars). Many of these applications have become integral parts of our lives, especially as we now see a wide range of devices being shipped with high definition cameras like smartphones, smart TVs, smartglasses, and drones. However, as we are embracing new technologies privacy concerns are also surfacing since these cameras create, process and transfer personally identifiable information to an extent that often remains unknown to those being affected by the technology. Therefore regulators are now investigating particular applications of computer vision [52] and there is a growing need for tools that inform users about what data is collected and what choices they have with respect to how the data is used.

In this paper, we focus on the use of facial recognition because this technology has not only improved in accuracy and performance that surpasses human performance in certain cases [52], but also seen wide spread adoption and steady growth in the commercial sector [4]. By definition facial recognition refers to a biometric technology that identifies individuals based on their distinctive and measurable facial patterns. Traditionally, facial recognition technology has been utilized by government and law enforcement agencies to support various security and safety operations [53, 7], but in recent years many commercial applications have started using facial recognition technology. As a result the U.S. Government Accountability Office (GAO) has provided a broader definition of facial recognition [52] that covers – (1) detecting faces in an image; (2) estimating demographic characteristics such as age, gender, race, nationality and religious belief; (3) determining facial expression and emotion; (4) verifying the identity claimed by a person; and (5) identifying an individual by matching an unknown image to a gallery of known people. Obviously, by extension, this very same technology can also be used to track people’s whereabouts, their activities, who they tend to hangout with, what items they spend time looking at in stores, whether they look healthy, and more. In this paper we refer to all the above functions as possible applications of facial recognition technology.

A report by Allied Market Research suggests that the global market of facial recognition technology is likely to grow to $9.6 billion dollars by 2022 [19]. The rise of Internet-of-Things (IoT) is in part responsible for this growth as the decreasing cost of cameras and computation devices have enabled large-scale deployments of IoT cameras in places such as schools, company workplaces, restau-
rants, shopping malls, and public streets \[23\]. As a consequence, we have started seeing real-world commercial applications of facial recognition technology in large-scale. Theme parks like Disney automatically group pictures into personalized albums for identified park users \[11\]; Facebook launched its facial-recognition based photo-sharing app Moments in the EU and Canada in 2016 \[18\]; and Retailers and advertisers have been using facial recognition technology in digital signage (e.g., smart TVs or kiosks) to deliver relevant ads to viewers based on their demographics information \[52\] \[21\]. Advertising companies are also looking into analyzing customers’ facial expression and sentiment towards different ads to increase attention and engagement of customers \[13\] \[2\] \[1\]. Other commercial applications that are likely to become more widespread in the near future include: automatic border control biometric kiosks \[14\], video-enabled ATM booths \[15\], smart vending machines \[12\] and automated menu suggester \[17\].

Widespread adoption of facial recognition technology poses an increasing threat to privacy. It can be used not only to identify individuals or track their whereabouts, but also figure out their social activities \[52\] such as with whom and where they hang out and their psychological state \[4\] such as whether they look depressed, tired or sick. Moreover, it is possible to combine facial data with auxiliary data (e.g., life style) to gain insights into people’s lives. Privacy advocates criticize the silent nature of facial recognition technology due to its lack of transparency of how video streams captured by cameras, at times concealed, are used \[56\]. One of the fundamental principles associated with information privacy is the right to “Notice and Choice”. However, current applications of facial recognition technology lack adequate/effective mechanisms for informing users of not only the presence of cameras but also the collection, usage, share and retention of sensitive data. In addition, a number of the described practices would ideally be required to provide choice options to users under certain regulatory bodies. But things only get worse as existing regulations on using facial recognition technology often fall short on recognizing the threats it poses. For example, in the U.S., no federal law explicitly regulates commercial uses of facial recognition technology, and existing laws \[6\] do not fully address the key privacy concerns that stakeholders have raised \[52\].

With this gap in mind we propose a privacy-aware infrastructure that not only notifies users of the existence of cameras nearby but also enables them to opt in or out of facial recognition systems. Our approach focuses on only those use cases of facial recognition technology where the use of facial recognition is optional, providing benefits to only those deploying it as well as those being monitored. For such use cases we improve the transparency of the systems and offer ways for users to control what data is collected about them. We hope this will lead to better transparency and higher acceptance of facial recognition technology.

2. Background and Related Work

2.1. Advances in Facial Recognition Technologies

Facial recognition has been an active field of research since early 1970s \[41\]. For many decades the progress in facial recognition was slow due to challenges arising from the fact that faces are not rigid objects, but are constantly changing due to aging, facial expression, makeup, or (facial-)hair style \[39\]. Jafri et al. provide a comprehensive survey of the evolution of facial recognition technology \[37\], including breakthrough systems such as Eigenfaces \[31\] and Fisherfaces \[27\].

Over the past decade, facial recognition technology has become faster and increasingly accurate in identifying individuals. In 2012, the adoption of deep convolutional neural networks (DNN) accelerated the attainable accuracy, and today Facebook’s DeepFace \[49\] and Google’s FaceNet \[47\] yield near-human accuracy. Unfortunately, these DNN-based systems are trained with private datasets containing millions of proprietary images from social media and are not robust against sub-optimal images. However, there have also been breakthroughs for reconstructing and identifying individuals from sub-optimal images. Savvides et al. have showcased new techniques for facial recognition that allow to reconstruct entire faces from partial images and even from the periocular region alone \[40\] and research by Foooprateepsiri et al. demonstrates the possibilities of creating 3D models of faces from 2D images \[34\]. In terms of facial expression current implementations are capable of tracking a person’s facial features in real-time under different facial expressions and face poses \[55\] \[29\] \[50\]. Such technical advances provide greater legitimacy to government and corporate entities’ argument for adopting facial recognition technology.

2.2. Privacy-Aware Image Recognition

Privacy-aware image recognition involves modifying certain contents, like faces, in an image or video stream to make such contents non-recognizable. Such image modification processes are often referred to as image denaturing. More complex forms of denaturing require identifying specific faces or landmarks within the image. The type of modifications that can be made is limited by the complexity, accuracy, and speed of image processing algorithms currently available, as well as the quality of the data and the details needed by the underlying service. Most of the commercial applications that we have previously discussed require face detection at real-time, but recognition and identification of faces are not necessary in all scenarios. Privacy-aware image recognition involves modifying certain contents, like faces, in an image or video stream to make such contents non-recognizable. Such image modification processes are often referred to as image denaturing. More complex forms of denaturing require identifying specific faces or landmarks within the image. The type of modifications that can be made is limited by the complexity, accuracy, and speed of image processing algorithms currently available, as well as the quality of the data and the details needed by the underlying service. Most of the commercial applications that we have previously discussed require face detection at real-time, but recognition and identification of faces are not necessary in all scenarios.

Privacy-aware video streaming has received attention by
researchers in recent years. Simoens et al. propose a denaturing approach that only works on a low-frequency video stream \[48\]. Aditya et al. provide a denaturing mechanism where specific faces can be obfuscated but their approach requires many seconds of processing per image and is thereby not usable for real-time applications \[25\]. Bo et al. present a reversible scheme for obfuscating faces \[28\]. However, their approach is not practical as it requires users to bear a printed bar-code in their clothes. Raval et al. propose a mechanism to block objects in images that have been either physically or virtually tagged by users at near realtime \[46\]. However, their approach depends on users manually tagging public regions in images. Jana et al. explore image transformations that preserve privacy but at the same time allow many perceptual applications to operate accurately \[38\]. In this work we primarily focus on designing an infrastructure that is capable of not only notifying the presence of nearby cameras but also enforcing real-time video denaturing at full frame rate by utilizing the face denaturing system proposed by Wang et al. \[54\]. Here, denaturing is achieved by replacing faces with black boxes, since other techniques like blurring are not resistant against machine learning attacks \[44\].

3. Regulations and Guidelines

**Regulations:** Facial recognition technology can be utilized for many commercial applications, but the extent of its current use is not fully known. To make things worse a report from the United States Government Accountability Office (GAO) noted that federal law does not expressly regulate how commercial applications use facial recognition technology \[52\]. There are a handful of laws regulating the collection and use of biometric identifiers, including facial images, in special situations such as when the data belongs to students or is used in connection with driver’s licenses \[16]\.

In the U.S. only two states, Texas \[10\] and Illinois \[9\], have adopted privacy laws that expressly regulates commercial uses of biometric identifiers, including scans of face geometry often used by facial recognition technologies. Both the Texas and Illinois laws require explicit consent from individuals before collecting any biometric identifier. They also prohibit sharing of biometric identifier with a third party, unless the disclosure meets a special exception. Furthermore, the law governs the retention of biometric records, including protection and destruction of biometric information after a certain period of time.

The upcoming General Data Protection Regulation (GDPR) in the European Union considers facial recognition as biometric data which falls in the same category as data about ethnicity, political or religious beliefs, and genetic data \[3\]. Therefore the law for legally collecting facial recognition data is higher than that of other types of personal data, and will require an explicit consent.

**Guidelines:** Several stakeholders including government agencies, industry trade organizations, and privacy advocacy organizations have proposed guidelines or best practices in using facial recognition technology commercially. Most of these guidelines often focus on the translation of Fair Information Practice Principles (FIPPs) into principles that are specific to facial recognition technology. Following is a summary of different guidelines proposed by International Biometrics & Identification Association \[35\], American Civil Liberties Union \[26\], Digital Signage Federation \[30\] and Federal Trade Commission \[32\]. Data collectors should:

- disclose in privacy policies what types of and for what purposes biometric data is collected;
- explicitly notify individuals when facial recognition is being used and obtain affirmative consent before using facial recognition to identify an individual;
- clearly notify individuals when using facial recognition is used to determine demographic characteristics;
- restrict or provide individuals with a choice to share any data collected through facial recognition technology with third parties;
- allow users to access, correct, and delete personal data;
- implement a specified retention period for personal data and dispose such data after the retention period expires.

In addition to best practices some stakeholders have advocated for “privacy by design” approach, which focuses on ensuring that a given technology is geared toward providing consumer privacy protections at every stage of the development. The privacy by design approach includes ideas such as encryption of facial recognition data, segregation of biometric data from other personal data, and automatic deletion of biometric data after a specified retention period \[52\].

**Our Goals:** Facial recognition based commercial applications have not yet become ubiquitous, but it is essential to address privacy concerns before it gains widespread adoption. At the very least consumers should be notified of sensitive data collections taking place around them and at times users should have the option to provide explicit consent in the form of opting-in or withdraw their consent in

---

2American Civil Liberties Union opted for a stricter guideline – forbidding facial recognition technology to determine an individual’s race, color, religion, sex, national origin, disability, or age.
the form of opting-out of services. This is more pertinent for facial recognition based technology because unlike other biometrics, such as fingerprint identification, facial recognition technology can be used to capture a face remotely and without the individual’s knowledge. Former chairwoman of the Federal Trade Commission (FTC), Edith Ramirez, has asserted that knowledge and consent are still the key components to protecting consumer privacy. However, providing substantive knowledge and meaningful consent is an acutely difficult challenge for informational privacy. As stated by Acquisti et al. people often do not know their own privacy preferences for a particular context, and the specific harms that may arise are not always tangible.

We try to address this problem by introducing an infrastructure that provides users with in situ notification and available privacy choices (e.g., opt-in/opt-out). This approach potentially enables users to understand the particular context under which their data is being collected, and thus enables them to make a better informed decision.

4. Users’ Privacy Concerns and Notification Preferences Regarding Facial Recognition

As part of a larger research project we conducted a vignette study to elicitate privacy preferences and comfort level of users towards a wide variety of data collection scenarios covering eight different factors – location, device type, data type, purpose, retention time, sharing of data, inference of data and the party benefiting from the data collection (in another study currently under review). Each factor can have different levels (e.g., location could be “at home” or “at work”), and after a manual selection of realistic scenarios the study used 380 scenarios of which 18 included the use of facial recognition on video data. An example of a scenario is: “You are at a coffee shop. This store has a facial recognition system that scans customer’s faces automatically as they enter the store in order to remotely identify returning customers. This method is used to keep track of your orders and make suggestions based on your ordering habits regardless of whether you are a member of their ‘coffee club’ or not. Your picture will be kept for a few hours.”

We had a total of 1007 participants on Amazon Mechanical Turk (50.1% female; average age 36.1 with a standard deviation of 10.9). Each of the participants was presented with at least one of the facial recognition scenarios. We analyzed the effects of the different factors with respect to the expressed comfort level towards the scenario, whether they would allow or deny the data collection as well as the participants’ interest in receiving notifications if they were to encounter such scenarios in their daily life.

We found that participant expressed a high discomfort towards scenarios that involved facial recognition and the collection of other biometric information like fingerprints or iris scans. 65% of the participants said that they were uncomfortable or very uncomfortable with a scenario involving facial recognition, e.g., a scenario for automatically checking out books from a library using facial recognition technology. Other data sources like presence sensors (33%) and temperature sensors (24%) raised much less concerns.

Since each scenario consisted of eight factors we used generalized linear mixed model regression with a random intercept per participant to measure the effect of different factors on the expressed comfort level. Biometric data related to facial recognition, iris scan, fingerprint had the most negative impact on participants’ comfort level, with facial recognition being the technology where participants felt most uncomfortable. At the same time participants did not completely reject the technology and were slightly comfortable with those practices they thought are happening today. Still, if asked whether they would allow or deny the collection of facial recognition or iris scan data, they would deny it (data on fingerprints was not significant with respect to this question). In general, answers varied depending on the specific scenarios hypothesized. For example participants were less resistant to facial recognition being used to identify customers in library versus departmental stores.

As mentioned above we also asked participants how often they want to be notified about a specific data collection. Although, participants expressed a variety of notification preferences, the regression model showed that in cases were biometric data was used and participants were not told about how the data was used (i.e., purpose was unspecific) their agreement to the statement – “I want to be notified about this data collection every time,” was significantly higher.

The results of our study show that users are skeptical about computer vision applications like facial recognition, iris and fingerprint scanners, although they might feel comfortable with the status quo of how the technology is used. More importantly we found that contextual factors such as the purpose of a data collection have a large influence on both the willingness to participate and their desire to be notified about data collection. These results underline the need for a flexible infrastructure that can help to inform users as well as enforce their privacy choices.

5. Privacy-aware Notification Infrastructure

To address the issues described in the previous section we designed and developed a privacy-aware notification infrastructure that not only notifies users of the existence of nearby sensors (e.g., camera) but also enables them to opt-in/opt-out of any service using sensitive data if applicable. In the following section we will first briefly describe the different components of our privacy-aware infrastructure, before illustrating how we integrated it with a face denaturing system to design a privacy-aware video streaming system.

---

2Comfort level was expressed in a likert scale of 5
Our infrastructure consists of the following components.

**Internet of Things Resource Registry (IRR):** The purpose of the IRR is to allow IoT resource owners (i.e., those that are responsible for the setup of cameras) to easily publish privacy-related information as well as advertise services that are built on top of such resources. An IoT resource can refer to apps, services, sensors, virtual sensors as well infrastructure elements that might be collecting and/or using user data. The IRR offers a web interface that can be used to specify privacy policies in a form that is compliant with a machine readable JSON schema. The schema represents information commonly found in privacy policies, e.g., on the web [45] or as required by law (e.g., FIPPs [20]). It comprises of fields to specify the responsible party, purpose of data collection, retention time, granularity of data collection and third-party data share. More importantly, resource owners can also specify control options for users to regulate how their data is used, if applicable.

The IRR guides resource owners through the process of creating a privacy policy by slicing the policy schema into separate parts, providing information about what might help users understand the policy and making sure all the required information is entered. Fig. 1 shows a screenshot of the different policy related information captured through the IRR.

The IRR is designed as a distributed platform that can be set up not only on building levels but also on a larger scale like cities. While the first may be controlled and managed by the organizations that own the building, the latter may be open to input from anybody that has created or knows of data collecting resources. The open scenario requires moderation of entries and anti-spam mechanisms that are not yet part of our setup. The availability of an IRR can be locally advertised with bluetooth beacons, or discovered through a central directory based on geo-location of the user.

**IoT Assistant (IoTA):** The IoT assistant can be thought of as a browser for IoT resources as it allows users to discover those resources and services in their vicinity (those that have been entered in an IRR). It is currently implemented as an Android application and automatically retrieves policies from the IRR relevant to the user’s current location. The IoTA lists resources and services available (see Fig. 2) and informs users about the each resource’s functionality, its owner, as well as what data it collects and what it does with that data (e.g., if data is shared, how long data is retained, if data is aggregated, etc). The service list gives an overview of the apps or websites that offer functionality based on the data collected by the resources.

In its current implementation, the IoTA lists all resources it discovers. Over time, we plan to develop filters that can be configured by a user, and eventually models users’ preferences to selectively decide when, how often, and what to show to the users. We also envision the IoTA to serve as a personalized privacy assistant (similar to a personalized privacy assistant for smartphone apps [42]). With time and usage the app will learn a user’s privacy preferences; notify her about mismatches between her preferences and the observed policies, and may also semi-automatically configure privacy settings on her behalf. Over time, we believe the availability of privacy settings to become more prevalent in different scenarios, in part because of emerging regulatory requirements, including, as discussed earlier, requirements to obtain opt-in consent from users in some situations.

![Figure 1: Screenshot of the IRR](image1)

![Figure 2: IoT resources and services available in the vicinity.](image2)

**Policy Enforcement Point (PEP):** The IoTA also allows users to configure privacy settings, if available (see Fig. 4a). Changes made in the IoTA are sent to a policy enforcement point (PEP) that ensures that the privacy settings are accurately captured and enforced. The set of privacy choices available depends on the underlying services that offer simple REST APIs to enforce the privacy settings. While it is possible to provide flexible privacy settings to users, for the purpose of this study we only provide simple out-in and opt-out options. However, our PEP is capable of offering temporal, spatial and contact based privacy choices to users.
Our PEP maintains a database for storing each user’s privacy settings, e.g., to disable facial recognition during specific times of the day or when one is at a specific location.

6. Privacy-aware Video Streaming

We integrate our privacy-aware notification infrastructure with a video denaturing system to build a privacy-aware video streaming service. Our proposed system informs users of nearby cameras when they approach the vicinity of deployed cameras. It also provides users with an opt-in choice (as our default policy is opt-out) to facial recognition based services. We have developed an automated class attendance app as one possible application. Other use cases where facial recognition technology best fits our infrastructure include automated menu suggestion, admission to transportation systems or checkout kiosks. In the following section we will first briefly describe the different components of the face denaturing system. Next, we will describe the interactions that take place among the different components. Lastly, we present some performance and scalability results. The face denaturing system proposed by Wang et al. [54] consists of a Face Trainer and a Privacy Mediator component.

**Face Trainer:** The face trainer uses OpenFace, an open source face recognition library that performs face recognition using deep neural networks (DNN) [22], to capture and extract the facial embedding of users who desire to opt-in (or depending on the application may choose to opt-out) to our facial recognition system. Currently, users send images of their face through an Android app. Users can upload as many image frames as they desire (the more the better), but users are instructed to upload at least 20 frames for training purpose. Users have to sign-in using Google’s OAuth to upload data to the training server. Users’ email addresses are used as identifiers to process setting requests made from the IoTA. Fig. 2(b) shows a screenshot of the app.

**Privacy Mediator:** The Privacy Mediator is the component responsible for denaturing live video feeds. Each camera has its own Privacy Mediator that is comprised of two main modules namely RTFace and Registration. RTFace performs the denaturing of images in real-time at full frame rate while the Registration module registers relevant privacy policies and practices with our IRR. Details regarding the type of DNN and image processing algorithms used by the Privacy Mediator are available at [54].

6.1. Work Flow Among All Components

The following steps describe a typical work flow of how a user interacts with our system. Fig. 3 pictorially illustrates the overall interactions.

1. We assume the user has installed the IoTA app on her smartphone and has discovered that cameras are being used to perform facial recognition in her vicinity. She can review the privacy policy associated with the technology and decide, e.g., on whether or not she wants to support the purpose for which the data is collected. The IoTA also shows to her that a service called ‘Automated Class Attendance’ (shown in Fig. 2b) is using facial recognition technology to register class attendance.

2. If the user decides to use the ‘Automated Class Attendance’ service she first has to opt-in using the IoTA as shown in Fig. 4a. Once the user makes her choice the corresponding setting is sent to the PEP for update. For the automated facial recognition to function she has to upload images of her face to the Face Trainer server by downloading the ‘Automated Class Attendance’ app where she authenticates herself with her Google account. She can then use the phone’s camera to upload images of her face to the Face Trainer server (as shown in Fig. 4b).

3. The PEP upon verifying (checking OAuth token) the request sent by the user, first updates the local database with the user’s current setting and then forwards the request to the Privacy Mediator for appropriate action.

4. Later on when the user enters the proximity of the camera (e.g., enters the class) the IoTA picks up the signal of the beacon attached with the camera and notifies the use of camera. The IoTA app can also potentially send a heartbeat signal to the Privacy Mediator to indicate that the user is near the viewing scope of the camera. This helps the Privacy Mediator narrow down the potential candidates currently visible in the system.

5. The Privacy Mediator, upon receiving the request from the PEP, first retrieves the user’s facial embedding from the training server and then starts performing facial recognition to detect the user’s face in the video stream. Depending on the privacy setting selected by the user the Privacy Mediator either denatures the user’s face or
retains unaltered video frames (a screenshot of the denatured video feed is shown in Fig. 5). If the user opts in the Privacy Mediator upon receiving images of the user entering the room can register her attendance (see Fig. 4).

Figure 4: User opting into Automated Class Attendance service. Face intentionally blocked for review purpose.

(a) IoTA Notification  (b) Face Training  (c) Attendance

Figure 5: A screenshot of how users’ faces are obfuscated where users have the capability to opt-in to revealing their faces using the IoTA app (as shown on the right side of the figure).

(a) Face Obfuscated  (b) Privacy Setting

6.2. Performance Analysis

In terms of performance the Privacy Mediator and PEP are the most critical two components in our infrastructure. In this section we therefore focus on analyzing the performance of these two components. We use the same network structure and parametric settings for the DNN as described by Wang et al. [54].

6.2.1 Accuracy and Scalability of Privacy Mediator

Denaturing faces from a large pool of users poses challenges to both the accuracy and scalability of the Privacy Mediator. In the context of a large-scale deployment, such as campus-wide or city-wide, the total number of individuals to be recognized can be considerably higher. However, even though the total user population may be large, the number of people captured by a camera at any given time is usually limited. Moreover, our IoTA can send heartbeat signals whenever it enters the purview of a camera (determined by the beacon signal strength) to reduce the number of potential candidates that the Privacy Mediator has to identify.

Wang et al. [54] evaluate their system by computing the time it takes to predict users when the pool of users varies from 10 to 100. They found that on average it takes less than 100 milliseconds to recognize a user from a pool of 100 users. The average frame rate for the denatured video feed was around 31.2 frames per second in processing two concurrent HD video streams (Privacy Mediator ran in a VM with an Intel 4-core i7-4790 processor @ 3.6 GHz). In terms of identification accuracy, the accuracy drop from 99.5% to 91.9% as the number of users is varied from 10 to 100.

One current limitation of the face denaturing system is that it is not bulletproof in the sense that few frames with a human-recognizable face can slip. However, Wang et al. have proposed different ways to reduce such privacy leaks [54]. Our goal is to not provide a bulletproof facial obfuscation system rather an infrastructure that supports discovery of nearby cameras and available privacy settings. We expect the accuracy and speed of facial recognition technology to improve in the near future; in which case many of the current limitations can be satisfactorily addressed.

6.2.2 Scalability of PEP

Considering that a single PEP may be supporting a large amount of users and a wide variety of services in the future, we need to analyze how well it scales as the number of users increases. Currently, our PEP supports the following three basic operations– a) status request: returns the current status of a user’s privacy setting; b) opt_in request: change a user’s current preference to opt_in; c) opt_out request: change a user’s current preference to opt_out.

Whenever a request is forwarded to the PEP it first verifies if the right user has sent the request. This verification step is done through OAuth tokens where the IoTA sends the user’s gmail address along an OAuth token to the PEP. The PEP then forwards the OAuth token to Google’s authentication sever. Google’s authentication sever responds back with a either a valid or invalid response. Depending on the response received from the Google authentication sever the PEP finally responds with either a ‘Successful’ or ‘Unsuccessful’ message back to the IoTA. Fig. 6 illustrates the overall work flow.

For each request the total service time can be divided into two parts: 1) Google authentication time; 2) database query and update time. We used SQLite database to store users’ settings. To capture the amount of time spent in each of these phases we record three timestamps. At the beginning of processing the request, we record time \( t_1 \). After authentication finishes, we record time \( t_2 \) and finally we record time \( t_3 \) right before sending the response back to IoTA. We can
then compute the following three operational time:\endnote{4}

time to authenticate Google token = $t_2 - t_1$

time to query and update database = $t_3 - t_2$

total time to process a request = $t_3 - t_1$

To compute these times we send out parallel status, opt_in and opt_out requests to the PEP\endnote{5} To emulate a large number of users we associate a random user_id with each incoming request to PEP. Each user_id generates a separate entry in the local database residing in PEP. We randomly generated 1000 user_ids. For the purpose of generating valid OAuth tokens we randomly associate each outgoing request from PEP to one of 10 valid gmail account. Table 1\endnote{1} shows our findings. We can see that as the number of concurrent requests increases, the average authentication time also increases significantly. Given that the request packets are very small in size (<1 KB) and the network bandwidth for PEP is in the order of 3000 Mbps, most likely the primary bottleneck is caused by the queuing time at Google’s authentication sever.\footnote{6} However, even with 100 concurrent requests the total time required was around one second (on a Intel 2-core Xeon, 2.50GHz processor). Also we have to keep in mind that users, most likely, are not going to be making very frequent updates to their privacy settings.

![Figure 6: Communication between IoTA and the PEP.](image)

Table 1: Scalability of our PEP

<table>
<thead>
<tr>
<th>Number of concurrent requests</th>
<th>Google authentication time (ms)</th>
<th>Database query/update time (ms)</th>
<th>Total time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>151.89</td>
<td>9.77</td>
<td>161.66</td>
</tr>
<tr>
<td>100</td>
<td>1106.24</td>
<td>8.30</td>
<td>1114.54</td>
</tr>
<tr>
<td>500</td>
<td>2904.37</td>
<td>9.08</td>
<td>2913.46</td>
</tr>
<tr>
<td>1000</td>
<td>3982.91</td>
<td>8.67</td>
<td>3991.58</td>
</tr>
<tr>
<td>2000</td>
<td>7201.44</td>
<td>8.85</td>
<td>7210.29</td>
</tr>
<tr>
<td>5000</td>
<td>11130.49</td>
<td>8.62</td>
<td>11139.11</td>
</tr>
<tr>
<td>10000</td>
<td>11478.99</td>
<td>9.68</td>
<td>11488.67</td>
</tr>
</tbody>
</table>

7. Discussion

Our infrastructure to notify and enforce user’s privacy choice can be generalized to many other IoT scenarios. As a matter of fact our infrastructure has already been deployed at UC Irvine and Carnegie Mellon University for advertising location-based services that track people’s whereabouts in campus through WiFi access points. In general our infrastructure is suitable for any scenario that involves data collection with sensors; in such scenarios our infrastructure helps to improve transparency by informing users of the different data collections taking place.

To foster adoption of our technology we are looking at ways to reduce deployment and maintenance efforts. We are, therefore, currently looking at ways to automate the setup of IRRs as well as the discovery of off-the-shelf IoT devices connected to the same network (e.g., through Manufacturer Usage Descriptions (MUD)).

We envision IoTA to act as a personalized privacy assistant as our infrastructure matures, and as more resources and services start enrolling into our system. We plan to leverage machine learning techniques to not only determine when to notify users of nearby sensors but also semi-automatically configure user’s privacy settings.

8. Conclusion

Legitimate concerns over privacy and misuse of sensitive data generated by facial recognition technology pose a significant obstacle to the widespread acceptance of such technology. If key privacy concerns can be satisfactorily addressed, enabling real-time analytics on video streams from public spaces can be valuable. In this context our proposed infrastructure can alleviate some of the key privacy concerns by providing users with in situ notification of nearby cameras and presenting users with privacy settings supported by the underlying system.

Our proposed infrastructure is applicable to a wide range of IoT scenarios; scenarios that involve notifying users of nearby IoT sensors. This is only the first version of the infrastructure; we are still working on automating many aspects of the infrastructure to make it more user friendly.

Acknowledgement

This research has been supported in part by the National Science Foundation under grant SBE-1513957 and CNS-1518865, as well as by DARPA and the Air Force Research Laboratory under agreement number FA8750-15-2-0277. The US Government is authorized to reproduce and distribute reprints for Governmental purposes not withstanding any copyright notation thereon. Additional support has also been provided by Google. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the Air Force Research Laboratory, the NSF, Google, or the US Government.

\footnote{4} We ignore the time it takes for a request to go from IoT to PEP and vice versa as this latency varies across WiFi or 4G connections.

\footnote{5} The gevent plugin \footnote{6} is used to boost the # of concurrent requests to 10,000

\footnote{6} Alternatively, Google could have treated such high volume requests from a single IP as a denial-of-service attack and hence the large delay.
References

[7] Street level surveillance. [https://www.eff.org/sls/tech/biometrics/faq]


