Enabling Privacy-Preserving First-Person Cameras using Low-Power Sensors

Muchen Wu, Parth H. Pathak, Prasant Mohapatra
Computer Science Department,
University of California, Davis, CA, USA.
Email: {muwu, phpathak, pmohapatra}@ucdavis.edu

***Abstract*—Wearable smart devices such as smart-glasses, smart-watches and life-logging devices are becoming increasingly popular, and majority of them are being equipped with first-person cameras. Such first-person cameras on smart-glasses or lifeloggers capture photos/videos from user's point of view, allowing them to record and share user's everyday events. However, these wearable devices with first-person cameras raise serious privacy concerns because they can also capture extremely private moments and sensitive information of the user. Currently, such devices lack the intelligence to understand user's preferences about certain scenarios being sensitive/private. To address this problem, we present PriFir, a scheme that enables *Pri*vacy-preserving *Fir*st-person cameras. PriFir is based on the idea that low-power sensors (e.g. accelerometer, light sensor, etc.) embedded in smartphones and smart-watches can be leveraged to identify sensitive scenarios. Learning from user's preferences, PriFir employs a cascade of classifiers that tags a scenario to be sensitive simply based on the characteristics of the low-power sensor data. We evaluate PriFir using real sensor traces spanning over multiple days and show that it performs highly accurate classification at a low energy cost.**

I. Introduction

There is a tremendous rise in wearable smart devices in last few years. New wearable devices such as smart-glasses [1], smart-watches [2] [3] and life-loggers [4] [5] are increasingly being adopted by end users. Many of the wearable devices are equipped with first-person cameras (also known as point-of-view cameras) which capture photos and videos from the first person's perspective. Examples of such devices include Google Glass [1] and life-logging devices like Narrative clip [4] and Autographer [5]. The life-logging devices can be configured to automatically take pictures every few seconds, allowing users to create long-lasting digital memories of their everyday events. First-person camera devices are also shown to be useful in health and safety (patients with memory loss and Alzheimer's [6]), personal informatics [7] and physical analytics [8].

While first-person camera devices are becoming ubiquitous, a major concern in their adoption is their privacy implications. Point-of-view photography can record the most private moments of the user. Common examples of such sensitive scenarios include the user visiting restroom, typing website password on computer, engaging in a private meeting, etc. Once captured, this imagery is vulnerable to the risks of being mistakenly shared by the user; or other malicious applications installed on such smart devices are able to leak the images. One report demonstrates a spyware on Google Glass called Malnotes [9] that deceptively acquires the permission to access Glass's camera and the Internet, takes photos every ten seconds and uploads the images to the remote server without user's awareness. Either way of unintentional image exposures may lead to serious public embarrassment probably with additional social and professional consequences. Moreover, with thousands of photos or hours of videos taken in one day via the first-person camera, it is unsurprisingly tedious and overwhelming for the user to manually inspect every one of them and check for any private information.

978-1-4673-7331-9/15/$31.00 © 2015 IEEE

Our initial survey about the design of such devices reveals that they lack the intelligence to understand user's preferences. The devices rely on the user to turn off the first-person camera when they are in sensitive scenarios. There exist some image processing based solutions [10] which analyze the image to locate pre-defined sensitive objects after it has been captured. The problem with such post-processing design is that the image has already been captured before being analyzed, making it vulnerable to the risks of unintentional sharing and stealthy leakage. Also, such image processing based approaches scale poorly on mobile devices, especially on energy-constrained wearables, due to their heavy computation load.

In this paper, we present PriFir, a scheme that enables *Pri*vacy-preserving *Fir*st-person cameras on wearable devices. PriFir is designed to identify *user-specific* sensitive scenarios purely using low-power sensors on smart devices. Because it entirely depends on whether the user considers a given scenario sensitive or not, PriFir needs to learn user's preferences with a small amount of training. At the core of PriFir is the idea that: first, low-power sensors such as accelerometer, light sensor, orientation, etc. can generate delicate signatures for various scenarios; second, with the combination of outstanding sensor features, PriFir identifies the sensitive scenarios with a very high accuracy. PriFir leverages various low-power sensors available in smartphone and smartwatch to determine a given scenario sensitive or not. If the scenario is tagged sensitive, all access requests to the first-person camera is denied to protect user's privacy.

There are three salient features of PriFir: (1) It performs the classification of a scenario being sensitive or not in advance, before a picture is taken (the camera is accessed). This eliminates the risk of the image being leaked before it

is analyzed for private information. (2) Because PriFir simply relies on monitoring of low-power inertial sensors, it is much more energy-efficient compared to image processing based techniques. (3) PriFir applies a cascade of machine learning classifiers to save energy, by making classification with lower-power sensors first and employing more energy-expensive classifiers later only if further analysis is necessary.

The contributions of our work are as follows:

1) We show that low-power sensors such as accelerometer, light sensor, orientation, etc. embedded in smartphones and smartwatches can be exploited to generate fingerprints of sensitive scenarios. In many cases, the features of only a few such sensors are enough to determine if the scenario is sensitive or not.
2) We design PriFir, a scheme that enables privacy-preserving first-person cameras purely relying on low-power sensors. PriFir consists of a cascade of low-complexity machine learning classifiers, each of which is trained using one or a few low-power sensors. As classifiers based on lower-power sensors appear first in PriFir cascade, most scenarios are classified early with low energy cost. PriFir is user-specific as it learns and adapts to user's preferences with small training period.
3) We implement PriFir and evaluate it using real sensor traces collected from smartphone and smartwatch. It is shown that PriFir can classify the sensitive scenario with an accuracy of 87% while restraining the false alarm rate (percentage of non-sensitive scenarios misclassified as sensitive) to be lower than 5%.

This paper is organized as follows. Section II described our privacy goals and approach. Section III shows how low-power sensors can be used to fingerprint sensitive scenarios. The design of machine learning classifiers and their energy-efficient cascade arrangement is presented in Section IV. Section V evaluates PriFir in terms of its accuracy and energy efficiency. We discuss related work in Section VI and conclude in Section VII.

II. Problem Description and Approach

In this section, we describe our problem using privacy goals, and provide the outline of our approach.

A. Privacy Goals

PriFir is useful in protecting user's privacy in the following cases.

A wearable first-person photography device (e.g. Google Glass, Autographer, Narrative clip etc.) is configured to take pictures periodically. User's primary concern when using such a device is that since it is configured to access the camera periodically, it can also capture certain personal and private moments which user would not like to share with others. Examples of such scenarios include when user is in restroom or in bedroom or when user is visiting her bank website on her laptop. Since the life-logging devices are likely to take hundreds of pictures in a day, it is extremely difficult for the user to scan each picture for privacy. In this case, our objective is to embed intelligence in the first-person camera device, which learns and determines when user is in a sensitive scenario and stops accessing the camera.

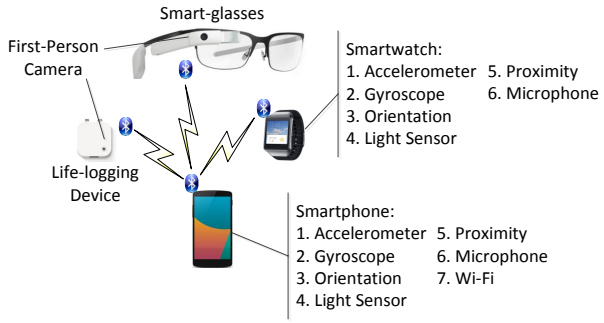


Fig. 1: Smart Devices and Available Sensors
(Images are taken from [1], [3], [4] and [12])

Apart from this, it is possible that one of user's smart devices is infected with a malicious software which can access the camera periodically. Such an application can take pictures from user's point-of-view and leak user's most private moments. [9] demonstrated a similar malicious app for Google Glass while [11] showed the same for the smartphones. In this case, our objective is to design an operating system service which determines if user's current scenario is sensitive or not, and sets the permissions to access camera hardware. We emphasize that our objective in this work is to protect user's privacy from her own first-person camera device. We leave the issue of protecting user's privacy from other user's camera devices to our future work (see Section VII).

B. Requirements and Challenges

We believe that our privacy-protecting scheme should also meet the following requirements along with the privacy goals described above. These requirements pose additional challenges which we describe here as well.

1) A Priori Decision Making: A crucial requirement is that a scenario should be tagged sensitive or non-sensitive before a picture is taken. This is different from other post-processing based approaches [10] which determines if the picture is sensitive or not after the picture is taken using image processing techniques. We believe that once the image is taken, there is a far greater risk of it being leaked even before it can be processed. A safer approach, on the other hand, is to proactively stop taking pictures when sensitive scenarios are detected, so we refrain from using image processing techniques in PriFir.

2) User-centric Model: Privacy is often subjective and user-dependent. The same scenario can be private to one user while being non-private to another. Our scheme should respect this user-centric phenomenon. It is challenging to design a scheme that can learn and adapt to user's preferences.

3) Scenario vs. Location: It is important to note that sensitive scenarios are not always sensitive locations. Activities like having a private conversation or typing personal information on a laptop can happen at multiple places, which apparently, location can not provide enough pattern to recognize. This means that determining user's contextual location (e.g.

bedroom, office etc.) is not sufficient and it is necessary to determine the scenarios which user finds sensitive.

4) Minimal User Intervention: A user would ideally like the device to make an informed decision without actively providing input. Our scheme should learn user's preferences and perform decision making independently afterwards.

C. Auxiliary Goal - Energy Efficiency

An important auxiliary goal while protecting user's privacy is to do so using lower energy consumption. Energy efficiency is necessary to be considered because all the devices in our system including smartphone, smartwatch and the life-logger are battery operated. Our scheme of determining if a scenario is sensitive or not should not consume excessive battery resources. For this reason, we use low-power sensors such as accelerometer, light etc. to fingerprint sensitive scenarios. This allows us to build a scheme which has lower cost of sensor data collection and lower processing complexity compared to a scheme where sample pictures are periodically taken and computationally expensive image processing is applied. We will quantify the energy efficiency of different sensors in Section IV.

D. Approach

Our presented solution, PriFir, is designed to meet the privacy goals and requirements mentioned above while achieving high energy efficiency. PriFir is based on the fact that today's users own multiple smart devices such as smartphone and smartwatch, each of which is equipped with plethora of low-power sensors. Fig. 1 shows variety of sensors available in latest smart devices. If we are able to exploit these sensors from both wrist and thigh areas in determining whether user's scenario is sensitive or not, we can achieve our privacy-preserving goals without requiring any computationally expensive methods such as image processing.

PriFir is designed in two steps. First, we show that low-power sensors can be used to classify a scenario in sensitive or non-sensitive with very high accuracy in Section III. Second, we note that most of the times one or a few low-power sensors are sufficient to achieve an accurate classification. We leverage this observation to build the PriFir machine learning classifier in Section IV. The classifier is a cascade of subclassifiers each of which is built using one or a few low-power sensors. Due to cascade arrangements, a given scenario is first tested with low-cost subclassifiers which are likely to result in accurate classification at a very low energy cost. PriFir operates as follows:

(1) PriFir collects data from low-power sensors of user's smartwatch and smartphone. During the initial training period, user actively indicates when she is in a sensitive scenario.

(2) With user's input and sensor values, PriFir builds a classifier as described above. This involves feature extraction, building subclassifiers and arranging them in order of energy cost to form a cascade.

(3) After the training, PriFir classifier observes the selected low-power sensors and classifies if user's current scenario is sensitive or not without requiring any user intervention.

TABLE I: List of low-energy sensor features, where *sp* and *sw* denote sensors in smartphone and smartwatch respectively. Features of Accelerometer, Gyroscope and Orientation are calculated for three axes. Refer to [14] for the feature definition and abbreviations.

Sensors: Features
Accelerometer-sp, Accelerometer-sw, Gyroscope-sp, Gyroscope-sw: DCMean, DCArea, ACAbsMean, ACAbsArea, ACEntropy, ACSkew, ACKur, ACQuartiles, ACVariance, ACAbsCV, ACIQR, ACRange, ACEnergy, ACBandEnergy, ACLowEnergy, ACModVigEnergy, ACPitch, ACDomFreqRatio, ACMCR, DCTotalMean, DCPostureDist, ACTotalAbsArea, ACTotalSVM
Light-sp, Light-sw, Sound-sp, Sound-sw, Orientation-sp, Orientation-sw, Proximity-sp, Proximity-sw: Minimum, Maximum, Mean, Median, Standard deviation, Skewness, Kurtosis

PriFir is built based on the assumption that user's smartphone and smartwatch are equipped with low power sensors such as accelerometer, gyroscope etc. This assumption is reasonable given that most current smartphones and smartwatches [13] are already equipped with these sensors. As shown in Fig. 1, the smartwatch and first-person camera device can communicate with the smartphone via Bluetooth LE. We refer to this communication network as a *wearable network*. The PriFir classifier can run on the smartphone which collects data from its own sensors as well as smartwatch's sensors. Once it has processed the data and made a decision, it informs the first-person camera device to set the camera access policy accordingly. The process is repeated periodically in order to constantly switch to new policy based on the current scenario.

III. Identifying Sensitive Scenarios

In this section, we provide the details of our data collection and the methodology. Note that our presented approach works for variety of scenarios specified by the user, and we pose three scenarios (i.e. typing, visiting restroom and being outdoors) as examples to show how they can be accurately classified using low-power sensors.

A. Data Collection and Methodology

Our objective is to exploit low-power sensors in smartphones and smartwatches to aid the first-person camera devices (e.g. smart-glasses or life-logging devices), as shown in Fig. 1. As commercial lifeloggers and smartwatches leave little flexibility to modify their design parameters, we employ three Google Nexus 5 smartphones [12] in our experiment for the ease of data collection. The first phone is hanging at the chest area to act as the life-logger, the second phone is placed at the wrist to operate as the smartwatch, and the third phone is set at thigh area as the personal smartphone.

For the smartphone and smart-watch devices, we collect the data for the following low-power sensors - Accelerometer, Gyroscope, Light, Proximity and Orientation. On both devices, AndroSensor [15] application gathers all sensor data 20 times per second. Acting as a life-logger, the smartphone hanging at chest area constantly records video through the first-person camera for the further ground truth analysis. One user carries all three devices for 3 days (16 effective hours in total), and continues with regular daily life at workplace, residence and other environment.

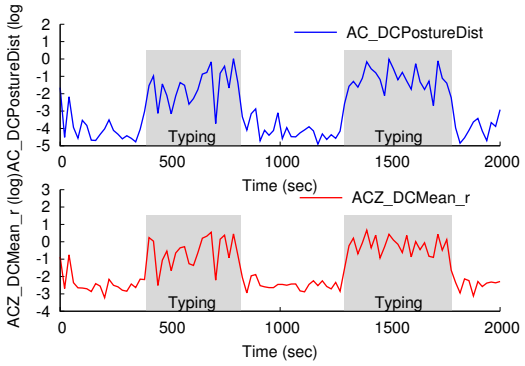


Fig. 2: Typing Recognition via Accelerometer Features

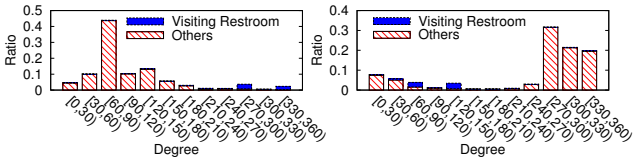
After data collection, we calculate 7 basic features for motion-irrelevant sensor data and 23 advanced features for the motion-related sensor data as described in Table I. Those 23 advanced features, presented in [14], are used to describe the characteristics of Accelerometer and Gyroscope, as they show strong correlation with plentiful human activities such as walking, standing, etc. Features are calculated for every piece of one-second-long sensor data. Note that this complete set of features only forms our preliminary set on which we apply the feature selection later. For the remaining sensors, we calculate 7 basic statistical features (i.e. min, max, mean, median, standard deviation, kurtosis and skewness). All features are calculated for sensors in the smartphone as well as the smartwatch. The video footage recorded from the life-logger is used to identify the time periods which the user regards as "sensitive" and "non-sensitive". As we discuss later, these identified time periods are used for analysis, training as well as the ground truth while testing PriFir.

B. Identifying Sensitive Scenarios using Low-cost Sensors

Based on the collected data, we now demonstrate how low-cost sensors in smartphone and smart-watch can identify sensitive scenarios. Here we choose three common example scenarios as representatives.

1) Typing: Typing is considered as a sensitive activity because a first-person camera can capture user's passwords or any other private information input via the keyboard. Here we are more interested in detecting typing activity when user inputs information on a laptop or a desktop computer, not within her wearable network (as defined in Section II). We claim that such typing can be detected using the accelerometer sensor in smartwatch. This is demonstrated in Fig. 2. It shows how two sample features of smartwatch accelerometer change when the user is typing. DCPostureDist calculates the differences between the mean values of the X-Y, Y-Z and X-Z axis and is shown to be indicative of sensor's (in this case hand's) orientation compared to the rest of the body posture [14]. Similarly, DCMean also shows clear variation when user is typing, because it is the static component of the acceleration which changes with body posture.

We build a machine learning classifier using 1-day data out of the entire dataset. The classifier uses simple logistic regression [16] with 10-fold cross-validation. From Table IIa, it can be observed that the low-power sensors such as accelerometer in smartwatch is sufficient to identify typing with True Positive (TP) Rate of 99%.



(a) Smartwatch (b) Smartphone

Fig. 3: Stacked Histograms of X-axis Orientation Mean Values for Two Devices

2) Visiting Restroom: We now consider a more complex sensitive scenario when the user is visiting the restroom. As discussed before, the problem can not be addressed using indoor localization in our case because of two reasons. First, there can be many cases where the user may or may not want the first-person camera to take pictures at the same location depending on the context. For example, the user might be willing to take pictures in bedroom during the afternoon but not at the night. Second, most indoor localization techniques depend on Wi-Fi signal fingerprinting which requires Wi-Fi scanning. The scanning is known to have a higher energy consumption [17], an undesirable effect for energy-restrained wearable devices.

By exploiting our dataset, the combination of many low-power sensor features can create a distinct signature of restroom visiting. The user visits 3 different restrooms during the experiment, i.e. one is at home and the other two are located in the working building. Compared to other activities, visiting restroom is a low-frequent event, comprising only 7.2% of effective experiment time. We examine through all low-power sensors and select two most distinguishing features as exemplars, the mean values of x-axis orientation for smartwatch and smartphone, presented in Fig. 3. Orientation features display strong correlation with the restroom visit. Compared to the rest occasions, the restroom visiting happens with a remarkably high percentage when the mean x-axis readings from smartwatch and smartphone fall into range $[270° - 300°)$ and $[120° - 150°)$ respectively. They appear in separate orientation ranges simply because two x axes of the devices point to two directions due to their different placement (i.e. wrist area and thigh area). Exploring the sensor features from different areas of body reveals the signatures of various scenarios. Presented in Table IIb, low-power sensors provide high TP Rate to distinguish user's visit to the restroom from other scenarios, which means that monitoring and analyzing low-power sensors can be used to instruct the first-person camera on and off.

3) Being Outdoors: When users are outdoors in public places, they commonly consider the scenario to be non-sensitive and would like the life-logging device to capture moments. In fact, capturing events in real-time at outdoors is the most attractive application of life-logging devices. Compared to the restroom scenario, distinguishing outdoor and indoor cases can be relatively easier due to the light

Class	TP Rate	FP Rate
Typing	0.998	0.011
Other	0.989	0.002

(a) Typing

Class	TP Rate	FP Rate
Restroom	0.784	0.045
Other	0.955	0.216

(b) Restroom

Class	TP Rate	FP Rate
Outdoor	0.948	0.058
Indoor	0.942	0.052

(c) Outdoor

TABLE II: Confusion Matrices of Logistic Regression in Three Scenarios
(True Positive Rate - TP Rate and False Positive Rate - FP Rate)

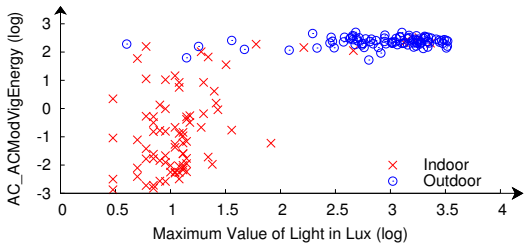


Fig. 4: Scatter Plot to Distinguish Indoor & Outdoor

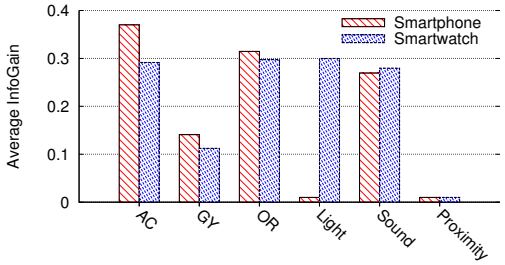


Fig. 5: Average Information Gain of Features for Each Sensor in Two Devices

sensor. Fig. 4 shows that the values of light sensor and smartphone accelerometer can clearly distinguish outdoors from the indoors. Apart from higher light sensor readings, the user is typically more active outdoors (e.g. walking, running, changing poses etc.) resulting in a noticeably different pattern of smartphone accelerometer. As shown in Fig. 4, ACModVigEnergy feature of smartphone accelerometer is able to capture high intensity activities. Table IIc shows the result of logistic regression based classifier that low-power sensors can distinguish outdoors from indoors nearly 95% of the time. The mis-classification is mostly due to some very high light sensor readings (for example when the smartwatch is placed right below the desk lamp) in indoor scenarios.

IV. PriFir Design

As demonstrated in the previous section, low-power sensors can be used to identify sensitive scenarios. Based on this, we present the design of PriFir in this section. PriFir can monitor and analyze the low-power sensors in smartphone and smartwatch, determine if the current scenario is sensitive or not, and allow or disallow the first-person camera to record images/videos accordingly. There are two steps in designing PriFir:

(1) We first evaluate different low-power sensors in terms of how well they can classify sensitive scenarios. At the same time, we also evaluate the actual energy consumption of monitoring these sensors in order to understand the average energy cost of classification using low-power sensors. Based on the accuracy and average energy cost of individual sensor, we choose a subset of low-power sensors that are useful in classification.

(2) We then develop multiple classifiers (referred as subclassifiers here onward) based on the selected set of sensors, and the subclassifiers are arranged in a cascade to form the PriFir classifier which performs classification in an energy-efficient manner.

A. Selecting Useful Subset of Sensors

If all features listed in Table I are used, it is possible to build a classifier that can classify sensitive scenarios with high accuracy. However, our objective is to perform such a classification using lower energy cost. We observe that not all low-power sensors are indicative of sensitive scenarios and hence it is possible to save energy by not monitoring them. To identify which sensors are not useful in classification, we first perform feature selection.

1) Information Gain-based Feature Selection: PriFir uses *Information Gain* [16] to evaluate the worth of a feature. Information gain is based on entropy of the provided data. Entropy is a measure of impurity in the provided data. Let D be the total number of training instances in which S are sensitive and N are non-sensitive. Entropy of D can be calculated as

$$Entropy(D) = -p_S \cdot \log_2 p_S - p_N \cdot \log_2 p_N \quad (1)$$

where p_S and p_N are the fraction of sensitive and non-sensitive instances. Information gain is the expected reduction in entropy when D is partitioned using a given attribute A. This way, information gain of an attribute A is calculated as

$$Gain(S, A) = Entropy(D) - \sum_{v \in V(A)} \frac{|D_v|}{|D|} Entropy(D_v) \quad (2)$$

where $V(A)$ is the set of all possible values of attribute A and D_v is the subset of D where the value of attribute A is v.

We calculate the information gain of all features of different low-power sensors for both smartphone and smartwatch. The information gain value is between 0 and 1 for any give feature. We rank all the features in descending order of their information gain and select the top 100 features. Fig. 5 shows the average information gain of such features for different sensors. Accelerometer and orientation sensors in both smartphone and smartwatch are likely to have more impact on classification. The light sensor on smartwatch is also found useful, though the smartphone light sensor is of little use because its value remains very low and invariant mostly when the phone is in the pant pocket. Sound sensors also show high information gain, but as discussed in the next section, the energy consumption of recording sound level is much higher compared to other low-power sensors. Other sensors such as proximity and gyroscope

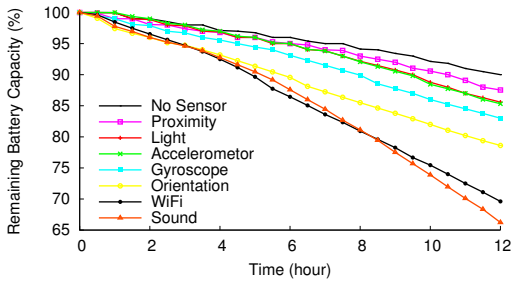


Fig. 6: Battery Capacity Depletion for Different Sensors

Sensor	Energy Cost	Sensor	Energy Cost
No Sensor	0.0	Gyroscope	0.6
Proximity	0.2	Orientation	0.9
Light	0.4	WiFi	1.8
Accelerometer	0.4	Sound	2.0

TABLE III: Energy Cost of Different Sensors

are found to have less usefulness in distinguishing between sensitive and non-sensitive scenarios.

2) Energy Depletion Rate of Monitoring Low-Power Sensors: We now take a detailed look at the energy consumption of monitoring different low-power sensors. To do this, we monitor each sensor individually for 12 hours with the sampling frequency of 20 times per second (20Hz), and measure the remaining battery capacity in percentage (%). As shown in Fig. 6, it is observed that accelerometer, gyroscope, orientation, light and proximity sensors have relatively lower power consumption. On the other hand, power consumption of monitoring sound level is much higher making it a less attractive choice for energy efficient classification. For comparison, we also provide the power consumption results for WiFi where the smartphone scans for surrounding WiFi access points and records their signal strength. As known from indoor localization works like [17], WiFi scanning consumes more power which is in line with our results.

We use the data presented in Fig. 6 to calculate the battery depletion rate for each sensor. This depletion rate is then used to calculate the *cost of monitoring* the sensor. Fig. 6 also shows the depletion rate of the smartphone battery when no sensor's data is being collected. This depletion is mostly attributed to running of Android operating system and other necessary background services. The cost of monitoring a sensor is expressed as the depletion in battery capacity per hour, and calculated simply by deducting the depletion rate when no sensor data is collected from the depletion rate of that sensor. The cost values are presented in Table III.

Generalized from Fig. 6 and Table III, gyroscope and proximity in both smartphone and smartwatch, as well as light sensor in smartphone, show little information gain and their rare usefulness. Then, sound and WiFi are proven to have higher energy depletion rates. Due to the benefit of high information gain and low energy cost, accelerometer and orientation sensors from both devices are selected to generate the possible subclassifiers, along with the light sensor in smartwatch. Hence, the further discussion regarding the proposed classifier is based on these five selected sensors.

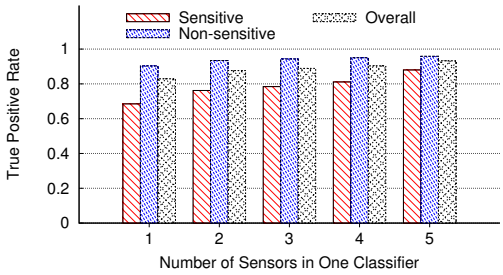


Fig. 7: True Positive Rates for Different Subclassifiers
(Corresponding Cost of Each Classifier: 0.4, 1.3, 1.7, 2.1 and 3.0)

B. Building PriFir Classifier - A Cascade of Subclassifiers

After PriFir finds a subset of five sensors that are useful for classification, it is possible to design many machine learning subclassifiers based on the selected sensors (individually and jointly). The challenge is to determine which subclassifier to use such that the classification accuracy remains high meanwhile the cost of the classification is low. There is a clear trade-off between the cost and accuracy when using different subclassifiers, especially with different number of sensors. Fig. 7 shows the TP rate of five different subclassifiers - each built using different subset of sensors. The subclassifier #1 is built only using the features of smartwatch accelerometer, while the subclassifier #5 is built based on features of five different sensors - accelerometer and orientation sensor in both smartphone and smartwatch as well as the light sensor in smartwatch. As more number of sensors are monitored and analyzed, their subclassifiers can achieve a better classification accuracy. Although one or a few sensors are often enough to classify between sensitive and non-sensitive scenarios, further improvement in classification can be obtained using additional data from other sensors. However, the cost of monitoring more number of sensors is also higher which shows the trade-off between energy cost and accuracy.

To address the cost-accuracy trade-off, we propose to use a cascade of subclassifiers as the PriFir classifier. Such a cascaded classifier has been previously applied for face and object detection [18], as well as spam email detection [19]. There are three steps in building the PriFir cascade:

(1) Multiple subclassifiers are built using all possible combinations of useful sensors. PriFir uses Logistic Regression to design the subclassifiers due to its simplicity and avoidance of any over-fitting. Once all subclassifiers are built, we use a TP rate threshold (TP_{th}) to remove the subclassifiers whose TP Rate for sensitive class is lower than TP_{th}. It ensures that the overall classification accuracy of the PriFir cascade is high, and enables us to create a tunable factor to manage the cost-accuracy trade-off. When TP_{th} is low, more low-cost subclassifiers will be allowed in the cascade which is likely to reduce the classification accuracy.

(2) The subclassifiers of useful sensors are first arranged in order of their cost. If two subclassifiers have the same cost, they are further arranged in descending order of accuracy. This mechanism ensures that a new scenario which should be tested

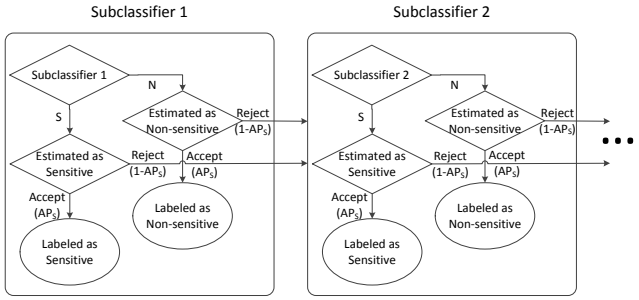


Fig. 8: Flow Chart of PriFir Cascade Model

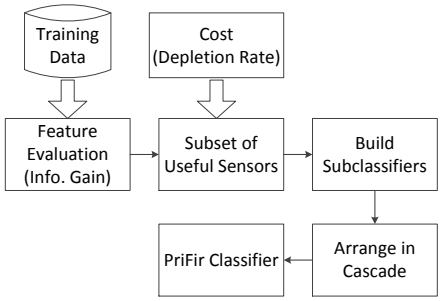


Fig. 9: Building Procedure of PriFir Cascade Model

whether it is sensitive (S) or non-sensitive (N) is first input to a lower-cost subclassifier.

(3) We derive the conditions for how a testing instance traverses through the cascade. The estimated class (S or N) of a subclassifier is accepted or rejected based on the acceptance probability (AP) of the output. This means that each subclassifier has two associated probabilities AP_S and AP_N for two estimated classes. Each one of APs determines corresponding estimation to be accepted or rejected. Here we show how AP_S is calculated, and the derivation of the AP_N follows similarly. AP_S can be calculated as

$$AP_S = P\{Actual\ Class\ is\ S | Labeled\ Class\ is\ S\} \quad (3)$$

Similarly, TP rate for sensitive class (TP_S) can be expressed as

$$TP_S = P\{Labeled\ Class\ is\ S | Actual\ Class\ is\ S\} \quad (4)$$

Let $P\{Act.S\}$ denote the probability that actual class of an instance scenario is S, and $P\{Lab.S\}$ denote the probability that the labeled class is S after estimation. $P\{Act.S\}$ and $P\{Lab.S\}$ can be calculated using the following

$$P\{Act.S\} = N_{ActS}/N_{Total} \quad (5)$$

$$P\{Lab.S\} = N_{LabS}/N_{Total} \quad (6)$$

where N_{Total} is the total number of training instances, and N_{ActS} and N_{LabS} are numbers of instances whose actual class is S and labeled class is S, respectively.

As TP_S can be obtained from each subclassifier, PriFir calculates AP_S for each subclassifier according to Bayes' theorem, with two probabilities in Equations (5) and (6).

$$AP_S \times P\{Lab.S\} = TP_S \times P\{Act.S\} \quad (7)$$

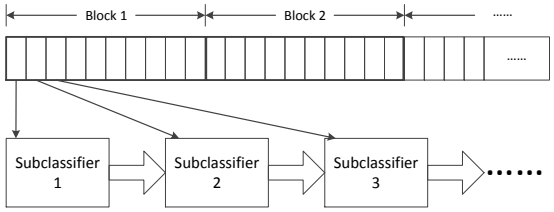


Fig. 10: PriFir Classifier during Testing Phase

$$AP_S = \frac{TP_S \times P\{Act.S\}}{P\{Lab.S\}} \quad (8)$$

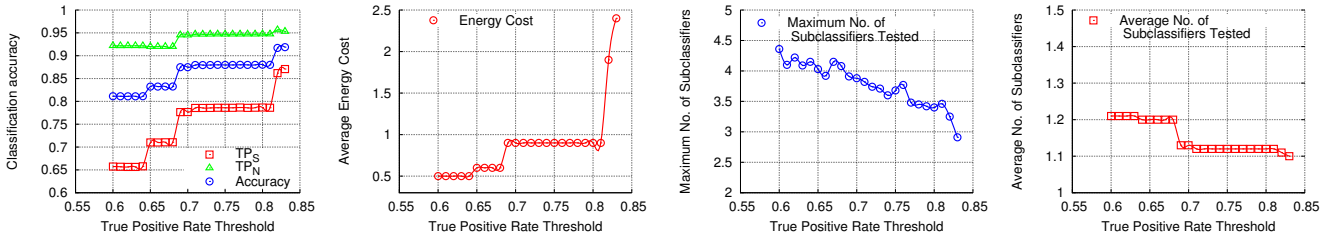
Fig. 8 presents detailed components in the model of PriFir cascade classifier. Fig. 9 shows the entire building procedure of PriFir cascade model. Note that the computational workload of building the PriFir classifier can be handled by current smartphones especially since the classifier is required to be built only once. However, this task can also be offloaded to user's other computer or cloud. In this case, once the cascade model is built, it can be imported back to the smartphone.

V. Evaluation

User-specific PriFir models are built up with various cascade structures corresponding to users' understanding of sensitive scenarios. Our PriFir model evaluation focus on the differentiation of user-defined sensitive scenarios from the rest regular non-sensitive events. In this section, one male user wears 3 smart devices and takes the 3-day experiment in three regular working days. Since one smartphone acting as the life-logger keeps the first-person camera recording all the time, the video footage (approximate 16 effective hours in total) is used to recognize the starting and ending time of one event, and serves as the ground truth for the evaluation. Scenarios like being in the restroom, typing on keyboards and staying in bedroom all belong to the sensitive category ruled by the participant. Dedicated to this particular example, we are able to discuss both accuracy and energy cost performance in details.

PriFir cascade is trained and tested using 3 days of data collected from user. We use first 2 days of data for training and building the PriFir classifier, and use the third day of data for testing. This reflects real-world behavior of users where a user proactively turns on/off the first-person camera device for first two days, marking as non-sensitive/sensitive scenarios respectively. PriFir uses this input along with the sensor data collected from the smartphone and smartwatch to train and create a classifier that is then capable of operating independently without user's intervention. We use 10-fold cross-validation on the training data to build the individual subclassifiers.

We use 10-second time blocks as shown in Fig. 10 which means that in the testing phase, PriFir determines the current scenario to be sensitive or not every 10 seconds. Once the decision is made, the camera access policy is fixed until the next decision. Specifically, PriFir starts the classification using first one-second sensor data tested with first subclassifier. It then accepts or rejects the output based on the acceptance probability of that subclassifier. If the estimation is accepted,



(a) Classification accuracy (b) Cascade energy cost (c) Maximum number of subclassifiers tested in the cascade (d) Average number of subclassifiers tested in the cascade

Fig. 11: PriFir's Performance with Different TP Rate Thresholds (TP_{th})

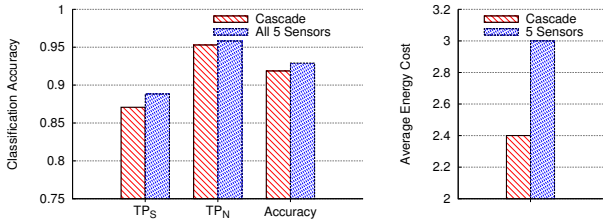
the entire 10-second block is deemed sensitive or not accordingly. If rejected, then sensor data is collected for the next second. Note that at any stage of the cascade, data is collected only for the sensors of that subclassifier. The procedure repeats until the given test scenario (10-second block) is classified as sensitive or not. We use Weka [16] to build the subclassifiers based on logistic regression, and then import them to our implementation of PriFir classifier. Note that we train and test the classifier offline after data collection. Also, we repeat the testing experiments 100 times to yield an average of classification accuracy and energy cost.

We first evaluate PriFir's performance with respect to the TP rate threshold (TP_{th}). Recall that TP_{th} is defined as the minimum TP rate of sensitive class that a subclassifier should have in order to be included in the cascade. It acts as a tunable parameter where its higher value achieves higher classification accuracy but also causes faster battery depletion (higher average energy cost). Figs. 11a and 11b show the classification accuracy and the average energy cost of the PriFir cascade for different values of TP_{th}. We can observe that both classification accuracy and average energy cost increase with the increase in TP_{th}. This is expected as more and more subclassifiers with lower energy cost (and lower accuracy) are eliminated from the cascade as TP_{th} increases. It is also observed from Fig. 11a that TP rate of sensitive remains lower than the TP rate of non-sensitive class. The accuracy in Fig. 11a is a weighted average of TP rates for both the classes. A sharp increase in classification accuracy and energy cost in Fig. 11a and 11b is observed at two TP_{th} values - 0.69 and 0.82. This is because at $TP_{th} = 0.69$, all subclassifiers based on only one sensor are excluded from the cascade since they can not meet the TP rate requirement. Similarly, at $TP_{th} = 0.82$, all subclassifiers with two sensors are excluded from the cascade, resulting in increase of energy cost as well as the accuracy.

Second, Figs. 11c and 11d show the worst and average case performance of PriFir cascade in terms of number of subclassifiers that needs to be tested before the given block can be classified. Both figures demonstrate that a classification can be obtained faster as the value of TP_{th} increases, because higher-accuracy subclassifiers appear early in the cascade. This tendency also matches with the trend of energy cost in Fig. 11b that faster classification also incurs more energy cost.

$\mathbf{TP_{th}}$	0.6	0.65	0.7	0.75	0.8	0.83
FP	7.81%	8.01%	5.49%	5.30%	5.29%	4.71%

TABLE IV: False Positive Rate: The Percentage of Non-sensitive Instances Misclassified as Sensitive



(a) Classification Accuracy (b) Average Energy Cost

Fig. 12: Performance Comparison between PriFir Cascade Model ($TP_{th} = 0.83$) and a Classifier using All 5 Sensors

Third, we now evaluate the FP rate of PriFir which is defined as the percentage of non-sensitive instances misclassified as sensitive. It reflects how much false alarm the model raises. Higher false alarm rate reduces the usability of the system from user's perspective because many non-sensitive scenarios where user would like to take pictures are misclassified as sensitive. Table IV shows how the FP rate varies with different values of TP_{th}. As we can observe, PriFir is able to achieve approximately 5% FP rate, indicating its high usability.

Forth, we compare the performance of classification accuracy and energy cost between PriFir cascade model and a classifier which uses all five sensors. In the latter, we do not use any cascade of subclassifiers but instead use all five sensors (accelerometer and orientation sensors in both smartphone and smartwatch and the light sensor in smartwatch) that were found to be useful in Section IV-A. We use Logistic Regression to build the classifier. Because such a classifier requires all five sensors to be monitored all the time, it allows us to evaluate how much energy savings are provided by the cascade arrangement. Fig. 12a shows the comparison of classification accuracy and Fig. 12b shows the comparison of energy cost. We observe that reduction of TP rate for sensitive class due to use of cascade arrangement is very low (around 2%) while the energy savings can achieve 25%. When $TP_{th} \geq 0.8$, the cascade arrangement incurs approximately 49% less energy cost compared to the classifier based on all 5 sensors.

VI. Related Work

The related work is categorized in following three topics.

(1) Wearable First-Person Devices: SenseCam [20] first presented a wearable first-person camera based device for life-logging as an aid for retrospective memory. Since then SenseCam has been adapted for numerous application in research including treating memory impairment [6], personal informatics [7], activity recognition [21], etc. Additionally, first-person cameras are becoming pervasive due to recent popularity of smart-glasses. Smart-glasses have enabled numerous new directions of research such as physical analytics [8] and attention-driven networking [22].

(2) Privacy in Wearables: Privacy issues related to wearable devices have gained significant attention from the research community recently. [23] did a thorough study about people's privacy management of life-loggers. The solution of protecting user's own privacy from these lifelogging devices was first addressed in PlaceAvoider [10]. Their proposed solution analyzes the images taken by the first-person camera and compares it with features of images that are previously classified by the user as private. Different from computationally expensive image processing used in PlaceAvoider, PriFir relies on low-power sensors only. There is also some recent research in protecting one's privacy from other user's camera. MarkIt [24] presents a visual privacy control system where a user can specify regions or objects that she want to protect from being captured by other cameras. Similarly, [25] presented a solution where a smart-glass can indicate privacy preferences to another wearable camera which is trying to capture video/pictures. Smart-glass uses Infrared LED (typically visible to RGB cameras but invisible to human eye) to communicate preferences about whether or not the user would like to be recorded.

(3) Continuous Sensing and Energy Efficiency: Energy efficiency is especially important when enabling new applications based on continuous vision, audio or context sensing. The energy minimization problem for continuous vision sensing was presented in [26]. Although the energy efficiency problem is similar to PriFir, their work is not concerned with the privacy issues. A recent work GlimpseData [27] presented a solution for continuous vision sensing where instead of detecting human faces in every frame of the captured video, low-power sensors are first used to predict if the frame will have a human face in it or not. Similar to PriFir, GlimpseData uses low-power sensors for higher energy efficiency, however it is not related privacy problems which are central to the design of PriFir.

VII. Conclusions

In this paper, we presented PriFir, a scheme that enables privacy-preserving first-person cameras. We showed that different scenarios can be categorized as sensitive or non-sensitive using the low-power sensors embedded in smartphone and smartwatch. We provide the procedure to build the PriFir classifier based on user's training. The PriFir classifier is a cascade of subclassifiers that can classify the sensitive scenarios with very low battery depletion cost. We evaluated PriFir using real sensor traces and showed that it achieved a classification accuracy of 87.1% for sensitive scenarios with less than 5% of false alarm rate. In our ongoing work, we are extending PriFir to design a scheme that can protect user's privacy from cameras of other users. This will allow a user to specify her preferences to other nearby recording cameras with low communication and energy cost.

References

[1] Google Glass, https://glass.google.com.
[2] Apple Watch, http://www.apple.com/watch/.
[3] Samsung Gear Live, http://www.samsung.com/global/microsite/gear/gearlive_design.html.
[4] Narrative Clip, http://getnarrative.com/.
[5] Autographer, http://www.autographer.com/.
[6] M. L. Lee and A. K. Dey, "Providing good memory cues for people with episodic memory impairment," in *ACM SIGACCESS*, 2007, pp. 131–138.
[7] I. Li, A. Dey, and J. Forlizzi, "A stage-based model of personal informatics systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 557–566.
[8] S. Rallapalli, A. Ganesan, K. Chintalapudi, V. N. Padmanabhan, and L. Qiu, "Enabling physical analytics in retail stores using smart glasses," in *ACM Mobicom*, 2014, pp. 115–126.
[9] Google Glass Spyware, http://www.forbes.com/sites/andygreenberg/2014/03/18/researchers-google-glass-spyware-sees-what-you-see/.
[10] R. Templeman, M. Korayem, D. Crandall, and A. Kapadia, "Placeavoider: Steering first-person cameras away from sensitive spaces," in *Network and Distributed System Security Symposium (NDSS)*, 2014.
[11] R. Templeman, Z. Rahman, D. Crandall, and A. Kapadia, "Placeraider: Virtual theft in physical spaces with smartphones," in *Network and Distributed System Security Symposium (NDSS)*, 2012.
[12] Google Nexus 5, http://www.google.com/nexus/5/.
[13] Android Wear, https://play.google.com/store/devices.
[14] E. Munguia Tapia, "Using machine learning for real-time activity recognition and estimation of energy expenditure," Ph.D. dissertation, Massachusetts Institute of Technology, 2008.
[15] AndroSensor, http://www.fivasim.com/androsensor.html.
[16] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
[17] Y. Gao, J. Niu, R. Zhou, and G. Xing, "Zifind: Exploiting cross-technology interference signatures for energy-efficient indoor localization," in *INFOCOM, 2013 IEEE*, April 2013, pp. 2940–2948.
[18] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *CVPR*, 2001.
[19] J. Pujara, H. Daumé, III, and L. Getoor, "Using classifier cascades for scalable e-mail classification," in *ACM CEAS '11*, 2011, pp. 55–63.
[20] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, "Sensecam: A retrospective memory aid," in *Springer UbiComp*, 2006, pp. 177–193.
[21] H. Pirsiavash and D. Ramanan, "Detecting activities of daily living in first-person camera views," in *IEEE CVPR*, June 2012, pp. 2847–2854.
[22] L. Zhang, X.-Y. Li, W. Huang, K. Liu, S. Zong, X. Jian, P. Feng, T. Jung, and Y. Liu, "It starts with igaze: Visual attention driven networking with smart glasses," in *ACM MobiCom '14*, 2014, pp. 91–102.
[23] R. Hoyle, R. Templeman, S. Armes, D. Anthony, D. Crandall, and A. Kapadia, "Privacy behaviors of lifeloggers using wearable cameras," in *ACM UbiComp '14*, 2014, pp. 571–582.
[24] N. Raval, L. Cox, A. Srivastava, A. Machanavajjhala, and K. Lebeck, "Markit: Privacy markers for protecting visual secrets," in *ACM UbiComp '14 Adjunct*, 2014, pp. 1289–1295.
[25] A. Ashok, V. Nguyen, M. Gruteser, N. Mandayam, W. Yuan, and K. Dana, "Do not share!: Invisible light beacons for signaling preferences to privacy-respecting cameras," in *ACM VLCS '14*, pp. 39–44.
[26] R. LiKamWa, B. Priyantha, M. Philipose, L. Zhong, and P. Bahl, "Energy characterization and optimization of image sensing toward continuous mobile vision," in *ACM Mobisys*, 2013, pp. 69–82.
[27] S. Han, R. Nandakumar, M. Philipose, A. Krishnamurthy, and D. Wetherall, "Glimpsedata: Towards continuous vision-based personal analytics," in *ACM WPA '14*, 2014, pp. 31–36.