Kai Malcolm
Guidance from Dr. Luca Bonomi
BMIF 7380 - Project Final Report
Wearable Sensor Data Privacy
Vanderbilt University
05 / 04 / 2022

Abstract

Mobile data from wearable devices such as smartphones or smartwatches presents high-resolution, fine-grained data detailing users' physiological attributes over time: such data are extremely diverse, and have been shown to be highly unique, leading to concerns for user privacy. Typical privacy methods like k-anonymization may fail in high-dimensional circumstances, rendering these traditional privacy protection methods unsuitable. By applying differential privacy-like solutions, the risk for re-identification from leaked wearable data is shown to be greatly reduced.

Introduction

The proliferation of wearable devices has only just begun: today, the wearables market sits at $116 billion and is expected to have a compounded annual growth rate of up to 19.48% each year until 2028 (Mehra, 2021). Furthermore, over 56% of Americans already own at least 1 wearable (McCarthy, 2021). As these data become more accessible, the constant monitoring from wearable sensors may benefit many healthcare applications: for example, wearable devices can be used for the remote monitoring of patients, reducing deaths from delayed medical intervention. Such high resolution data also plays a critical role in the development of models that go beyond monitoring and extend into prediction: by measuring personalized physiological baselines, it may be possible to achieve early detection of chronic conditions. However, there are significant privacy concerns in the use of mobile data in real-world applications. High resolution mobile data spanning many dimensions raises the concern that users may be able to be uniquely identified, despite the removal of PHI from the data. While privacy mechanisms such as k-anonymity, L-diversity, and t-closeness could be used to protect user data, they are not without their shortcomings. In particular, k-anonymity, the most popular, struggles to provide privacy protection in the presence of an informed adversary. Additionally, the potential high uniqueness in the mobile wearable device data could lead to overly-generalized sanitized data, leading to poor usability.

In this project, we study the privacy risks for sharing mobile wearable device data. In our application setting, a research center collects mobile wearable device data paired with some clinical features (e.g., COVID-19 diagnosis) before being de-identified (removal of PHI) and made available for research tasks (see Figure 1). An informed adversary may use background knowledge (e.g., compromised user mobile profile data such as attributes like calories and steps posted to social media, data breaches, or hacks of wearable device data storage providers) to attack de-identified shared data and violate the privacy of data contributors. In mobile app data breaches, not only will user wearable data be released, but also their demographic information such as name and address, which increases the severity of this privacy risk. This project illustrates the privacy risks for a small dataset of FitBit activities, where we assess the adversarial capabilities with different amounts of background knowledge (e.g., increasing number of longitudinal activities) in terms of data uniqueness and accuracy for a distance-based attack. Our application could extend from common wearables today such as smartphones to future wearables (such as smart clothing or augmented reality headsets), which may gather EEG or EMG data and exacerbate the privacy risks explored herein (Landau, 2020). The main contributions for this project are:

1. We proposed two empirical privacy measures to assess the privacy risk in sharing mobile wearable device data. First, inspired by De Montjoye, et al., we developed a uniqueness score, which quantified the uniqueness of the individual users in our dataset. Second, we designed a distance-based attack, which we used to empirically evaluate the ability of an informed adversary to match compromised profiles with the shared data.
2. We designed a randomized response approach, which sanitized the data by perturbing the recorded activities. For this method, we assessed both the effectiveness in mitigating privacy risks as well as retaining the usefulness of the original dataset.
3. We conducted several evaluations on an open-source, real-world dataset comprising mobile activities from the FitBit of 33 individuals recorded over a month (moibus, 2020).

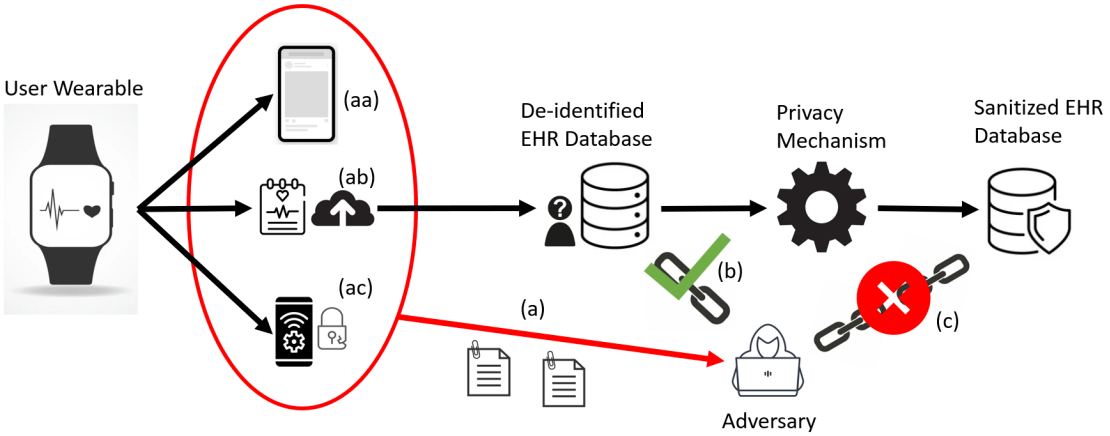

Figure 1: (a) compromised user profiles of wearable data may be obtained by an adversary via (aa) social media posts, (ab) intercepting users' data uploads, or (ac) profile disclosure due to data breaches of the apps or storage providers. In the case that wearable data is complemented by sensitive diagnoses and uploaded to a public research database, an adversary could leverage the compromised profiles to link and re-identify targets within the public de-identified dataset, but with an appropriate privacy mechanism in place, such linkage attacks may be prevented.

Methods

Privacy Risk and Uniqueness Assessment

Measuring the uniqueness of mobile data is challenging, as there are multiple attributes that could be used in combination to link users. To address this challenge, we measured the relative information content of each attribute by implementing KL-divergence (also commonly known as relative entropy), computed for every combination of attributes, generating a square, symmetrical matrix quantifying the relative information contained between each pair. From the said matrix, we select the attributes that maximize the average distance between individuals (see Figure 8 in appendix): from our evaluation, we select the "FairlyActiveMinutes", as this attribute well separates the individuals. Once this was conducted, the uniqueness is computed as the fraction of individual users in the data that can be uniquely identified by a subsequence of k-readings. In our evaluations, we generated an increasing number of readings (e.g., vectors), which are then matched with the original data. The uniqueness is measured on the basis of how many of those matches were the only match for the given user vector (e.g., that user's vector of data was found to be uniquely identified), similarly to (De Montjoye, 2013).

Distance-based Attack for Privacy Risk Evaluation

In order to further quantify the privacy risk, a distance-based attack was developed. The attacker uses the Euclidean distance to match a compromised vector of a target individual with the vectors in the shared data. Choosing a threshold value, the adversary determines that the target contributed to the data if the target can be matched to a vector within the shared data within a distance smaller than the given threshold value. We measure the success of such an attack in terms of accuracy, where higher accuracy indicates higher success in learning the membership.

Privacy Solution Design, Evaluation, and Data Utility Categorization

Given the small dataset size, adding any meaningful amount of noise in order to distort the data and achieve differential privacy would greatly compromise the data utility: instead, the chosen privacy solution was to implement a block randomized response, as inspired by Ding, et. al., 2020. This sanitization method creates a given number of bins to compartmentalize all the data, where the privacy parameter epsilon and the number of bins chosen both affect the data utility as well as the privacy protection (e.g. how well the adversary is able to successfully link back to the sanitized dataset). To partition the domain into bins, we use the noisy max technique to satisfy differential privacy. To find the noisy max, a random amount of Laplace noise (magnitude as a function of epsilon) was added to every value in the dataset, and the max of these perturbed data points was taken. The index of this perturbed max obtained the original (pre-noise) value: the noisy max is thus defined as that original value plus a new random value of Laplace noise (Dwork, 2014).

This process is visualized in the leftmost column of Figure 2 below, as well as the domain partitioning (binning) process described above (center column Figure 2). The final rightmost column of Figure 2 displays how the sanitization method functions. Our approach produces a sanitized activity by randomizing the original activity input with probability pB the algorithm returns a sanitized activity value from the same block as the input, else, it samples the sanitized values of the remaining bins. In our approach, we define pB as a function of epsilon and block size to ensure indistinguishability, with the goal to use the blocks to compress the domain size and facilitate the return of values within the same block to benefit data usability.

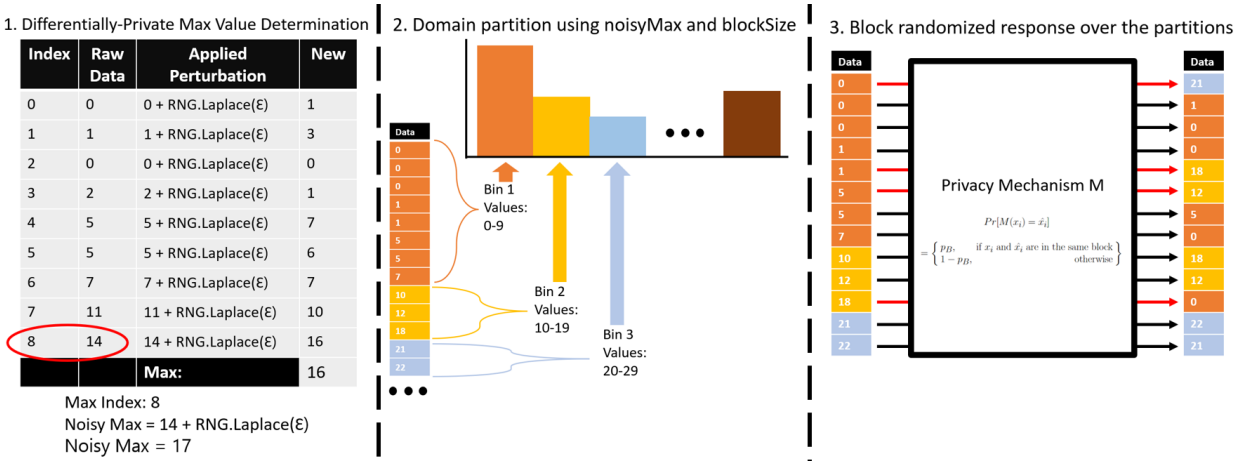

Figure 2: [Left] Visualization of the applied Laplacian noise in order to determine the noisy max in a differentially-private-like way. [Center] Domain partitioning using the number of bins and the noisy max, in order to discretize the dataset and reduce the domain. [Right] Applied sanitization method, showing how there is a random chance that some values are sampled outside of their respective home bin.

Results

In Figure 3 below, the uniqueness of the dataset illustrates that when matching directly against the raw data, just 3 readings can uniquely identify nearly 90% of all users. Clearly, as the number of readings increase, the percent uniquely identified increases. Note that increasing the threshold lowers the uniqueness, as would be expected when matching directly against the raw data: if matching the de-identified vectors back against the sanitized dataset, there should be a nonzero threshold which performs best.

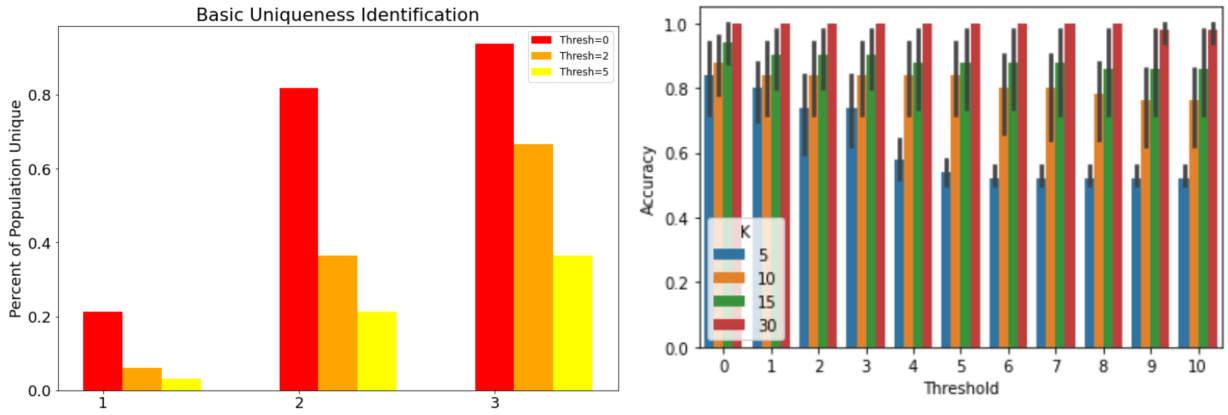

Figure 3: [Left] Percent of users in the database that can be uniquely identified using k (x axis) readings. Figure 4: [Right] Adversary accuracy when directly linking against the raw dataset, e.g. the dataset from which the attack vectors came from.

Figure 4 shows the accuracy of the adversary when linking the compromised user profiles back to the raw, de-identified data: while there is a clear trend that more readings increase accuracy, the main takeaway should be that even with just 5 readings, the accuracy is over 80%. Thus, Figure 5 and 6 below show the accuracy of an adversary when testing against our sanitized dataset, and shows how accuracy changes as a function of the threshold used, the privacy parameter epsilon, and K (the number of readings used). Note that in Figure 5, the accuracy of increased thresholds plateau after about 6 for larger block sizes, whereas for smaller block sizes, the accuracy monotonically decreases as threshold increases. Comparing Figures 5 and 6 (epsilon of 5 and 10, respectively), notice that as we relax the privacy protection, the adversary increases in accuracy, as one would expect. In particular, note that the accuracy achieved is around 50% on average: this is a significant reduction from attacking the raw dataset as found in Figure 4. Finally, from Figure 7 below, it is clear that while higher epsilons result in increased adversary accuracy, they also correspond quite strongly to minimize data distortions.

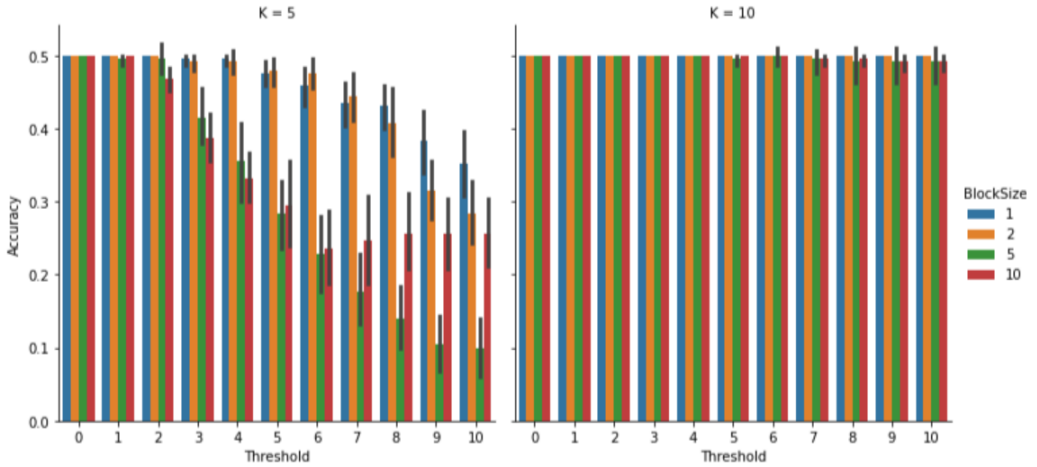

Figure 5: Adversary accuracy when linking back to sanitized data, where the sanitization method applied was based on an epsilon value of 5.

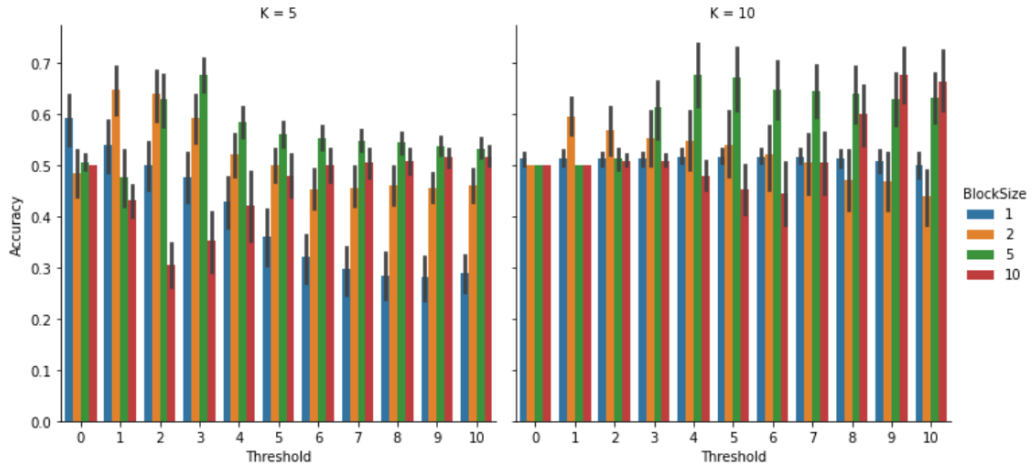

Figure 6: Adversary accuracy when linking back to sanitized data, where the sanitization method applied was based on an epsilon value of 10.

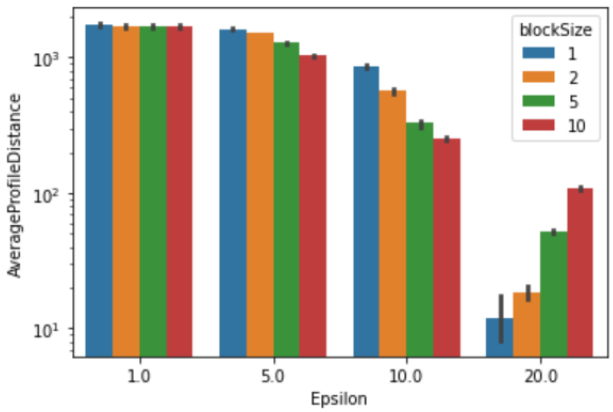

Figure 7: Data utility as measured in the amount of change in the average profile distance, shown as a function of the privacy parameter.

Limitations

A large constraint on this study was the small size of the dataset (just 31 users), which hampers any claims for generalizability of the results found. The other limitation pertains to the fact that, to date, there is no evidence for direct linkage attacks using wearable data, as this study attempts to prevent. However, the industry is still in its infancy, and other research in similar areas (GPS,

other “wearable” data from smartphones) suggests this could be a problem as seen in De Montjoye (2013) and Hern (2018).

Conclusions

This analysis successfully illustrated the high uniqueness of mobile data, and generated a distance-based attack algorithm which can be used to estimate the effectiveness of an adversary attempting to link compromised user profiles to released datasets, for both de-identified and sanitized datasets. Specifically, with just 3 data points, 90% of the individuals in this dataset were uniquely identifiable (explicitly: uniquely separable from the rest of the vectors in the database, not necessarily able to be uniquely linked back to their name and other PHI, as that was neither the aim of this project). Given the small dataset available, it is difficult to show the preservation of data utility, as even small amounts of noise greatly distort the dataset, constraining this analysis to using only large values of epsilon in order to preserve the data utility. As an example, with an epsilon value of 10 and a block size of 5, over 200 minutes of activity on average per individual are lost. For larger epsilons (more relative privacy protection), the smaller the block size, the better the data is preserved. Finally, a clear positive trend between the block size and attacker accuracy was established: despite this, the proposed sanitization method was able to almost halve the success rate of an adversary by introducing extra false positives and false negatives due to the data perturbation.

A last note is that all attributes are correlated: for an active lifestyle, one would expect higher values across the board for steps, calories, so on and so forth. An improved analysis could be finding the correlation between each column and creating a synthetic dataset that represents the fields with the highest correlations and span most of the other attributes.

Appendix

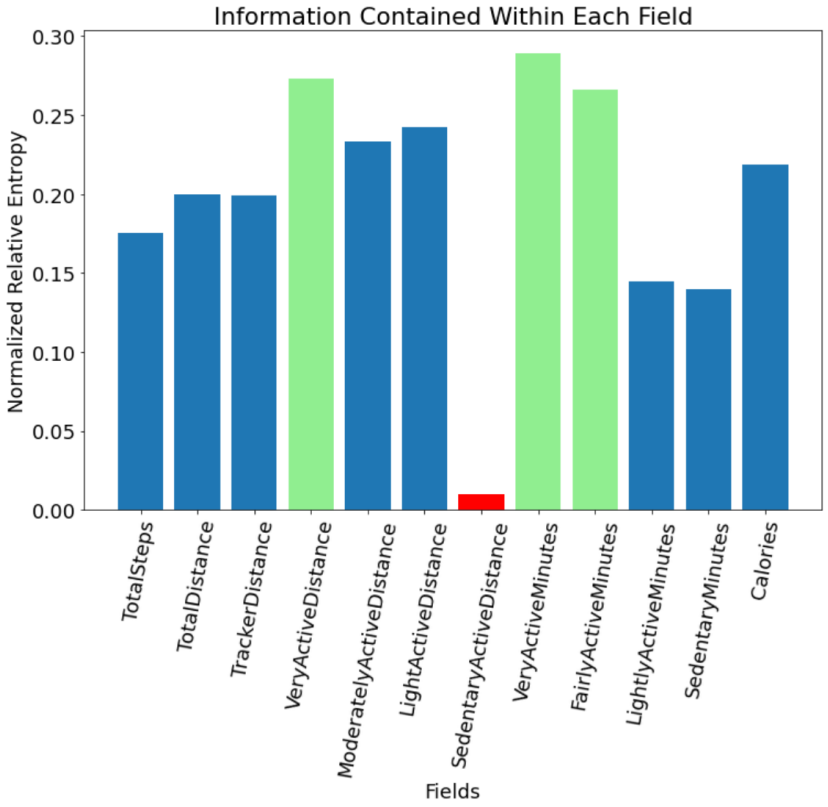

Figure 8: Relative information content within each attribute.

Privacy Analysis for Individual Activity:

Here, we study the privacy guarantee of our sanitization method M at activity-level. Given any two input activities x_1 and x_2 and any output $\hat{x}$, we want to bound:

$$\frac{Pr[M(x_1)=\hat{x}]}{Pr[M(x_2)=\hat{x}]}$$

Let $|B|$ be the size of the block in our equal block partition, an $|D|$ be the domain size, then:

$$\frac{Pr[M(x_1)=\hat{x}]}{Pr[M(x_2)=\hat{x}]} = \frac{pB/|B|}{(1-pB)(1/(|D|-|B|)} \leq e^{\varepsilon} \Leftrightarrow pB \leq \frac{e^{\varepsilon}}{e^{\varepsilon}+\frac{|D|}{|B|}-1}$$

References

Alam, M. (2021). *PRI-attack: Person re-identification attack on ... - arxiv.org*. PRI-Attack: Person Re-identification Attack on Wearable Sensing. Retrieved from https://arxiv.org/pdf/2106.11900v1.pdf

De Montjoye, Yves-Alexandre, et al. "Unique in the crowd: The privacy bounds of human mobility." Scientific reports 3.1 (2013): 1-5.

Ding, B., Kulkarni, J. (J., & Yekhanin, S. (2020, October 23). "Collecting telemetry data privately". Microsoft Research. https://www.microsoft.com/en-us/research/publication/collecting-telemetry-data-privately/

Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." Found. Trends Theor. Comput. Sci. 9.3-4 (2014): 211-407.

Hern, A. (2018, January 28). *Fitness tracking app Strava gives away location of secret US Army Bases*. The Guardian. Retrieved from https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Landau, O., Cohen, A., Gordon, S., & Nissim, N. (2020, April 21). Mind your privacy: Privacy leakage through BCI applications using machine learning methods. Retrieved from https://www.sciencedirect.com/science/article/pii/S0950705120302641

McCarthy, J. (2021, November 20). *One in five U.S. adults use health apps, Wearable Trackers*. Gallup.com. Retrieved from https://news.gallup.com/poll/269096/one-five-adults-health-apps-wearable-trackers.aspx

Mehra, A. (2021, April 15). *Wearable technology market worth $265.4 billion by 2026 - exclusive report by MarketsandMarkets™*. Wearable Technology Market worth $265.4 billion by 2026 - Exclusive Report by MarketsandMarkets™. Retrieved from https://www.prnewswire.com/news-releases/wearable-technology-market-worth-265-4-billion-by-2026--exclusive-report-by-marketsandmarkets-301269737.html

moibus. (2020, December 16). *Fitbit Fitness Tracker Data*. Kaggle. Retrieved from https://www.kaggle.com/datasets/arashnic/fitbit

Mordor Intelligence. (2021). *Smart wearable market trends, Size: Industry Growth (2022 - 2027)*. Smart Wearable Market Trends, Size | Industry Growth (2022 - 2027). Retrieved from https://www.mordorintelligence.com/industry-reports/smart-wearables-market.