What your Fitbit Says about You: De-anonymizing Users in Lifelogging Datasets*

Andrei Kazlouski[1,2], Thomas Marchioro[1,2] and Evangelos Markatos[1,2]
[1]*Computer Science Department, University of Crete, Greece*
[2]*Institute of Computer Science, Foundation for Research and Technology Hellas, Greece*

Keywords: Privacy, De-anonymization, Data Inference, Wearable Devices.

Abstract: Recently, there has been a significant surge of lifelogging experiments, where the activity of few participants is monitored for a number of days through fitness trackers. Data from such experiments can be aggregated in datasets and released to the research community. To protect the privacy of the participants, fitness datasets are typically anonymized by removing personal identifiers such as names, e-mail addresses, etc. However, although seemingly correct, such straightforward approaches are not sufficient. In this paper we demonstrate how an adversary can still de-anonymize individuals in lifelogging datasets. We show that users' privacy can be compromised by two approaches: (i) through the inference of physical parameters such as gender, height, and weight; and/or (ii) via the daily routine of participants. Both methods rely solely on fitness data such as steps, burned calories, and covered distance to obtain insights on the users in the dataset. We train several inference models, and leverage them to de-anonymize users in public lifelogging datasets. Between our two approaches we achieve 93.5% re-identification rate of participants. Furthermore, we reach 100% success rate for people with highly distinct physical attributes (e.g., very tall, overweight, etc.).

1 INTRODUCTION

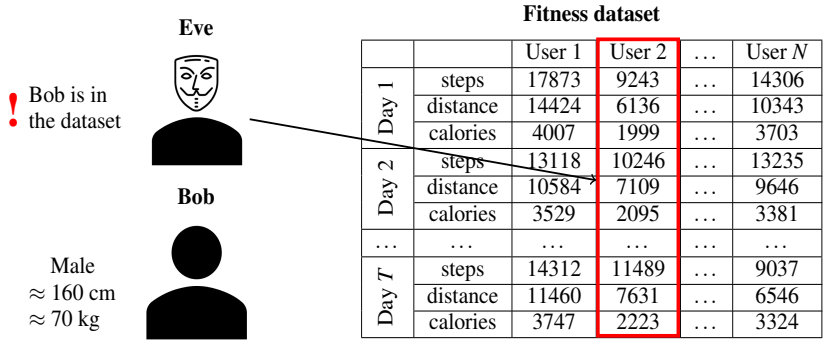
Smart watches and wearable fitness trackers have been gaining increasing popularity over the last decade. The current pandemic does not seem to have stopped fitness enthusiasts from purchasing devices to monitor their daily exercise. On the contrary, 2020 has seen strong growth in the market of wearables, which is forecasted to subsist also in 2022 and beyond (CCS Insight, 2021). Such devices are endowed with a number of sensors, and are able to measure a wide variety of fitness parameters, including steps, calories, sleep patterns, and in some cases even mood and stress levels. These data enable users to continuously monitor their progress, and adjust training schedule. Given the importance and sensitivity of the collected information, a number of privacy concerns have been raised in regard to extensive collection of fitness information.

The general consensus is that such ubiquitous data collection can lead to the production of a so-called quantified self, i.e., a state where a person is distinctly defined by their activity records. In this work, we investigate whether fitness records actually contain traces of the individual who produces them. More specifically, we try to de-anonymize users based on their fitness-related activities and/or their daily routine.

Our work is mainly focused on data from lifelogging experiments. Such experiments are aimed at collecting fitness records from participants who use wearable trackers through the day. Since these studies are often quite demanding for the entrants, the collected datasets tend to comprise a small number users. Therefore, a particular emphasis needs to be placed on protecting such participants' privacy. Typically, the revealing participants' details undergo a "sanitization" process where personal information is suppressed. In this paper we show that this might not be sufficient to fully protect the privacy of participants. We investigate whether an attacker – called Eve for the reader's convenience – can re-identify a target user – called Bob – in a public lifelogging dataset, us-

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813162. The content of this paper reflects the views only of their author (s). The European Commission/ Research Executive Agency are not responsible for any use that may be made of the information it contains.

Kazlouski, A., Marchioro, T. and Markatos, E.
What your Fitbit Says about You: De-anonymizing Users in Lifelogging Datasets.
DOI: 10.5220/0011268600003283
In *Proceedings of the 19th International Conference on Security and Cryptography (SECRYPT 2022)*, pages 341-348
ISBN: 978-989-758-590-6; ISSN: 2184-7711



Fitness dataset

		User 1	User 2	...	User N
Day 1	steps	17873	9243	...	14306
	distance	14424	6136	...	10343
	calories	4007	1999	...	3703
Day 2	steps	13118	10246	...	13235
	distance	10584	7109	...	9646
	calories	3529	2095	...	3381
...	...	...	...	...	...
Day T	steps	14312	11489	...	9037
	distance	11460	7631	...	6546
	calories	3747	2223	...	3324

Figure 1: In our first threat model, the adversary aims to link a person known to be in an aggregated dataset back to his fitness records. Assuming that the attacker has learned a basic profile of the victim, she infers physical attributes for all the users based on the daily fitness data, and chooses the "closest".

ing solely information she infers from the fitness data (i.e., without relying on personal identifiers). That is, she gains insight into Bob based *exclusively* on his fitness records and no other data. We assume that Eve is aware of her target's presence in the dataset, and that Bob's personal attributes have been removed. Under these assumptions, we study two approaches for Eve to re-identify Bob in the anonymized dataset.

De-anonymization based on Physical Parameters. The first method is based on inferring physical parameters from fitness data, and comparing the obtained results with the real-world information. In particular, we investigate whether it is feasible to deduce the user's gender[1], and whether they are overweight based on the Body Mass Index (BMI) margins. Since, given normal body type, people with BMI>25 are considered overweight, we choose the same threshold for our experiments. Furthermore, we study the possibility of identifying people who are taller than 177.6 cm, which is the average male height in Europe (World, er 1). Henceforth, when we say "overweight" we imply people whose BMI is higher than 25, and when we say tall/taller or short/shorter people we indicate people who are above or below 177.6 cm of height. To extract this information from daily records, we gather a number of open-source Fitbit data collections, and train cross-dataset inference machine learning models, using (i) daily steps, (ii) distance, and (iii) calories as features. Adopting a relatively low number of features increases the usability of our models, and allows us to visualize the obtained decision regions. Once Eve learns the physical characteristics of all the users in the dataset, she compares them with Bob's. If there is a single user with the same combination of parameters as Bob's, she concludes that such user is her target. If there are multiple, say k, the best she can do is choosing with probability $1/k$. Figure 1 depicts this threat model in more details.

De-anonymization based on Daily Routine. The second approach consists of de-anonymizing users based on their daily fitness patterns. Unlike the first method, the adversary does not need to know in advance the physical attributes of the target. However, it is necessary for Eve to be in possession of additional other target's data samples to re-identify them. Such extra samples might be obtained via social networks or through the target's medical records. For Fitbit, for example, it is possible to follow the fitness progress of friends in the dedicated app. Furthermore, a significant number of Fitbit users are members of the dedicated fitness communities where they share their progress with the world. The final predictions are based on the time series of fitness data. That is, we utilize a time series of length 24, where each entry represents an hourly tuple (from 00:00 to 23:00), containing: (i) number of steps, (ii) covered distance, (iii) burned calories, and (iv) the average heart rate during that hour. We couple these data with additional information on the day of the week, distinguishing between weekdays and weekends. Such distinction was made to account for possible changes in routine on Saturdays/Sundays. We train an LSTM-based inference model, achieving 93.5% de-anonymizing accuracy. This model essentially distinguishes users based on their daily routine, and times of the day when they are the most active. The biggest difference between our threat models is that the first one de-anonymizes users based on *"Who they are"*, while the second one categorizes individuals based on *"What they do"*.

The contributions of this paper can be summarized as follows:

- We train several inference models to infer physical parameters, and assess their performance.

[1]In this work by gender we imply the binary choice of male/female offered by Fitbit.

- We de-anonymize people based on fitness records, assuming their data are present in an aggregated dataset and their approximate physical parameters are known to the adversary. To execute this attack, we utilize previously obtained inference models, re-identifying minority[2] individuals in a lifelogging dataset. For the majority users, the adversary is still able to significantly reduce anonymity sets.
- We show that it is feasible to de-anonymize individuals in an aggregated dataset when the adversary possesses some records of their fitness data.

2 RELATED WORK

Previous works have studied the risks of ubiquitous IoT data collection and inference of private information from fitness data.

Inference from Fitness Data. The vast majority of the studies that investigated inference of personal characteristics from activity information relied mainly on raw sensor data (accelerometer, gyroscope, etc.) (Sathyanarayana et al., 2016; Parate et al., 2014; Dong et al., 2012; Kelly et al., 2017). Malekzadeh et al. proposed neural network-based approaches for anonymizing raw data produced by sensors (Malekzadeh et al., 2018; Malekzadeh et al., 2019). It is worth noting that these state-of-the-art works on inference from sensor data utilized small datasets (24 participants) likewise.
Some papers, however, explored the possibility of learning sensitive information directly from humanly understandable fitness data such as steps, burned calories, and covered distance. Torre et al. (Torre et al., 2018) investigated the correlation of the parameters. They also proposed a framework for privacy protection of fitness data. A re-identification attack on the fitness time series was introduced in a previous work of ours (Marchioro et al., 2021). Our former paper explored a similar threat model, where an adversary leverages daily fitness samples as a fingerprint for Fitbit users, re-identifying them with almost 80% accuracy. In this paper, we extend our work to more fine-grained samples that are available in some lifelogging datasets, achieving even better accuracy.

Ubiquitous Data Collection. A number of works have studied the privacy aspects of fitness trackers. The necessity to protect the Fitbit data and avoid building a "quantified self" was discussed in (Christovich, 2016). Hilts et al. (Hilts et al., 2016) did a comparative evaluation of wearables security, privacy and data sharing. The concern of users for sharing their fitness data was studied in (Vitak et al., 2018).

[2]We call "minority" individuals the participants whose characteristics are non-dominant in our datasets.

3 METHODOLOGY

This section describes the data used for the experiment, how they are leveraged for de-anonymization, and the architectures employed in our models.

3.1 Datasets

We gathered 3 open-source Fitbit datasets from various data sharing online platforms. All these datasets contain Fitbit fitness data in the same format. Table 1 depicts the statistics for the studied datasets. We distinguish overweight users based on the BMI threshold of 25, where $\text{BMI} = \text{weight}/\text{height}^2$.

Dataset D1. The Openhumans dataset (OpenHumans, 2016) comes from an online data sharing platform Open Humans. D1 contains weight and height of participants, and, hence, BMI. For the gender inference part of our experiment, we reconstructed the attribute from the names of the users, and dropped participants who have the unisex ones. Finally, we discarded users who do not have any recorded data.

Dataset D2. This dataset was generated via Amazon Mechanical Turk crowdsourcing marketplace (Furberg et al., 2016). D2 includes only participants' weight and height and has no record of their gender. We discarded all the empty (0 daily steps) entries, and dropped users for whom we could not compute BMI.

Dataset D3. The PMData dataset (Thambawita et al., 2020) was created during a 5-months lifelogging experiment, and counts 16 users who had been using the Fitbit Versa 2 wristband. D3 includes gender, height, and weight of all but 1 participant (for him only weight is missing). Unlike D1 and D2, this dataset was produced during a control experiment. Therefore, it has a similar amount of data for every user. After discarding the empty time series, the number of recorded days per participant is ranging from 80 to 152 days. Since, unlike D1 and D2, D3 is a proper lifelogging dataset (as defined in Sec. 1) we employ it as a benchmark for our de-anonmization approaches. We utilize D1 and D2 *only* for training the inference models for our first threat model.

Table 1: The statistics for the employed datasets.

Dataset	Total Samples	Males	Females	Overweight	Not Overweight	Tall	Short	Samples per User
D1	39225	18	13	15	16	9	24	17 - 3509
D2	480	-	-	9	4	4	9	2 - 49
D3	2119	13	3	7	8	12	4	80 - 152

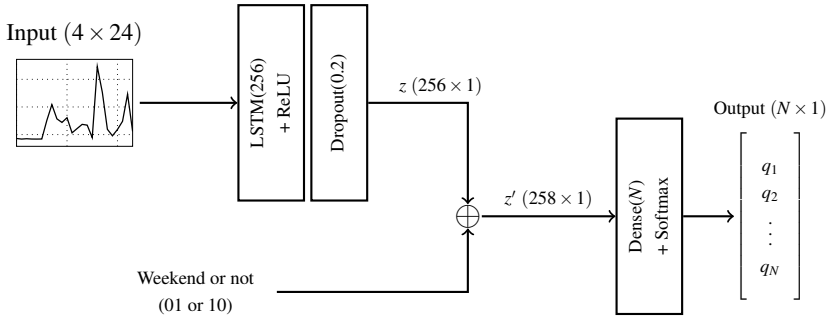


Figure 2: Architecture for routine-based inference. Two bits are concatenated to the output of the LSTM layer to model weekdays (01) or weekends (10).

3.2 Models and Data Utilization

Below we describe the employed approaches, and the training procedure.

Inference of Physical Parameters. We train 2-layers fully connected vanilla neural networks with early stopping. We employ the following hyperparameters for training:

- architecture: 120 hidden neurons + ReLU; 60 hidden neurons + ReLU; 2 output neurons + Softmax
- loss: binary cross-entropy, batch size: 64, optimizer: Adam, learning rate = 0.001

Training Procedure. For gender inference, since D2 does not include gender as a ground truth parameter, we utilize only D1 and D3: D1 for training/validation and D3 for testing. For the overweight and tall people detection, we train/validate our models on the combination of the D1 and D2 datasets, and test them on D3. We employ a 80/20 training/validation split for all the models. Moreover, we perform a 5-fold cross-validation, and choose the best model. We apply the trained models on D3 for user de-anonymization.

Routine-based User Inference. We train an LSTM-based neural network that takes a day of the week as an additional categorical input, distinguishing between working days, and weekends. More specifically, two bits – either 01 for weekdays or 10 for weekends – are concatenated to the output of the LSTM as depicted in Figure 2. We employ the following hyperparameters for training:

- architecture: as in Figure 2.
- loss: categorical cross-entropy, batch size: 64, optimizer: Adam, learning rate = 0.001

Training Procedure. Since D3 is the only lifelogging dataset, and has the most even spread of samples between all the users we employ it for our experiment (re-identifying users based on their daily routine). We utilize a 80/20 training/validation split. Moreover we perform a 5-fold stratified cross-validation.

3.3 De-anonymization based on Physical Attributes

In order to de-anonymize users from their physical parameters, we proceed in two steps: we first predict physical parameters from fitness records, and then search for users with unique tuples of parameters, who are identifiable.

User-wise Attribute Prediction. The models presented throughout this part of the paper infer binary information from single daily samples. That is, they are all maps from the feature domain $\mathcal{X}$ (i.e., all the possible combinations of steps, calories and distance) to $\{0,1\}$. In particular, we learn three binary maps: $\hat{q}_{\text{gender}}$, to estimate if a user is male or female; $\hat{q}_{\text{bmi}}$, to estimate if a user is overweight or not; $\hat{q}_{\text{height}}$, to estimate if a user is taller or shorter than 177.6 cm. However, if a user produced T samples, the map may yield different predictions for them. Thus, distinct predictions are combined to obtain a single, more accurate, prediction. The final prediction $\hat{r}$ is made according to a majority rule, which in a binary-decision setting

can be formalized as

$$\hat{r} = \begin{cases} 1, & \text{if } \frac{1}{T}\sum_{t=1}^{T} \hat{q}(x^{(t)}) > \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

meaning that if more than 50% of the samples are predicted to be positive, the final prediction is positive; otherwise it is negative.

User De-anonymization. Before any prediction is made, all the users $\theta_1, \ldots, \theta_N$ in a dataset belong to a same *anonymity group* of size N, meaning that an adversary can guess the correct user with probability $1/N$. We leverage the prediction models to answer the three binary queries, where each query splits a group into 2 subgroups. Therefore, three queries divide the dataset into $2^3 = 8$ anonymity groups, which may vary in size depending on the population of the dataset. It follows that, for a dataset with more than 8 participants, it is impossible to uniquely de-anonymize all of them. Nonetheless, if the target is a minority individual, the relative subgroup might be a singleton, making him/her easy to re-identify.

3.4 De-anonymization based on Daily Routine

The model that de-anonymizes an individual based on their daily activity predict one of the users in the dataset from single daily samples. The accurate map for this case is: $\hat{q} : \mathcal{X} \to \Theta$, where $\mathcal{X}$ is the domain of the features (i.e., all the possible combinations of time series for hourly steps, calories, distance, and average heart rate), $\Theta = \{\theta_1, \ldots, \theta_N\}$ is the set of the users present in the dataset, which has cardinality N.

4 RESULTS

Here we summarize the results that we obtain for the two previously introduced threat models, and illustrate our main findings.

4.1 De-anonymization based on Physical Parameters

Table 2 depicts the inference results for gender, overweight and height detection. In this table we report the accuracy our models achieve on the validation split, and the D3 test set that has not been observed at any point during training. We also describe the number of correctly classified users in the test split. We count a user to be classified correctly if the model is able to accurately predict the majority, i.e., $> 50\%$ of the samples for that person (Equation 1). We consider this metric as *the most important*, since it essentially illustrates the number of users, we are able to successfully infer physical characteristics of. Moreover, we report the respective f1 scores for each of the label for all the problems.

It is evident that all the models perform considerably better on the task of gender inference, achieving higher user and sample classification accuracies. This might be attributed to the fact that for the test dataset both BMI and height are non-binary parameters. Hence, it is significantly more challenging to classify users whose physical attributes are close to the classification thresholds. Overall, it is evident that the quality and quantity of the training data are enough to perform the accurate inference of physical parameters for previously unseen users.

Inference Visualization. In this section we display a graphical representation of our results, and discuss them. Since our models utilize only 3 features (steps, distance, and calories), we are able to construct 3D plots representing the decision regions defined by each model, i.e., the labels that are predicted for many possible values of the features. In order to make the regions visible, we evaluate rectangular grids of steps and distance for different fixed calories values. This way, we obtain the "layered" regions that can be observed in Figure 3 for the gender model where the layers are evaluated every 250 calories.

According to the Harris-Benedict (HB) equations (Harris and Benedict, 1918), daily basal calories (i.e., calories consumed just by basic metabolic functions,

Table 2: The adversary is able to infer physical attributes solely from the fitness records. Test accuracy is calculated over all samples, while user accuracy shows whether the majority of the data samples for each test user is classified correctly.

Attribute	Val acc	Test acc	Labels	Precision	Recall	F1	Users	User acc
Gender	0.925	0.925	Male	0.94	0.97	0.96	13/13	**1.000**
			Female	0.83	0.73	0.78	3/3	
BMI	0.81	0.731	Overweight	0.69	0.8	0.75	7/7	**1.000**
			not Overweight	0.65	0.71	0.71	8/8	
Height	0.968	0.821	Tall	0.87	0.89	0.88	12/12	**0.938**
			Short	0.67	0.64	0.66	3/4	

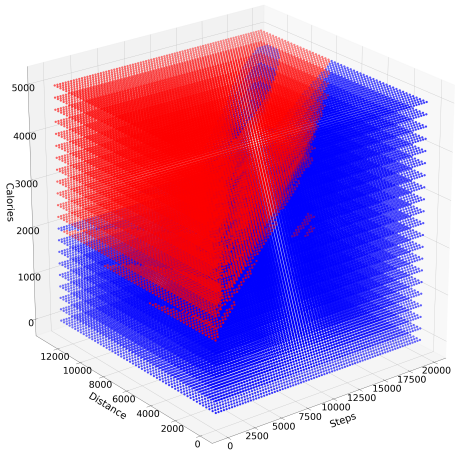


Figure 3: Decision regions for gender inference: blue color corresponds to females, and red to males.

without doing any exercise) burned by females and males can be estimated by two different empirical formulas. We illustrate those basal calories, and their Fitbit counterparts for all the users of the test D3 dataset in Figure 4. We can observe that all 3 females in the dataset burn considerably less calories in comparison to their male counterparts. We can verify that the trained model for gender inference (Figure 3) adhere to that pattern. Indeed, there are very few male samples below the 1500 calories plane, which seems to be fully in accordance with the HB equations, and Fitbit data. Furthermore, it can be observed that the areas with the same number of calories, but higher number of steps/distance ratio generally correspond to the female users. It can be interpreted that females - as they are typically lighter - need to take more steps/distance in order to burn the same amount of calories as males. These observations suggest that the obtained gender model's predictions reflect common sense. Likewise, the behaviour of the BMI and height models do seem to support reasonable assumptions, e.g., overweight people burn more calories than their counterparts given equal activity.

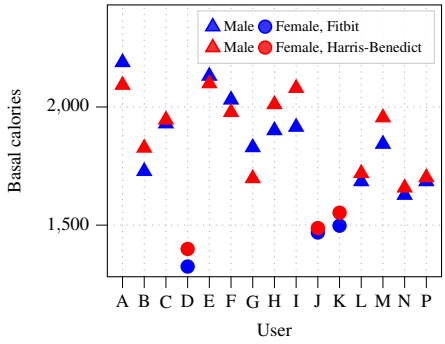


Figure 4: Daily basal calories for the test users that are estimated via Harris-Benedict equation and Fitbit.

De-anonymization from Predicted Parameters. If Eve knows that her target is present in a (not anonymized) lifelogging datasets, she can find them by looking at the personal characteristics of participants. That is why countermeasures such as k-anonymity (Samarati, 2001; Sweeney, 2002) and ℓ-diversity (Machanavajjhala et al., 2007) are often applied to public microdata. In this section, however, we show that even if personal characteristics are concealed, Eve is still able to re-identify some participants. In particular, she can leverage the queries introduced previously ($\hat{q}_{\text{gender}}$, $\hat{q}_{\text{height}}$ and $\hat{q}_{\text{bmi}}$) to detect "singular" participants.

Assuming that Eve is able to get the correct results for all queries, she is, thus, able to learn partial information on the participants of a dataset. The results for the the test dataset (D3) are displayed in Table 3. Therein, for each tuple of possible values for queries, we also display the number k of occurrences. That number indicates how many users share such combination of results. For an attacker, the most interesting tuples are, indeed, those with only one occurrence, because they correspond to a unique participant within the dataset. For instance, if Eve is searching D3 looking for a tall female with BMI<25, only one participant satisfies all the requirements (the last highlighted row in the table). Overall, for D3 the adversary is able to de-anonymize 3 minority participants based on their attributes with 100% probability.

Table 3: Occurrences (#) of query results for the test dataset. Those who are found with probability 1 are highlighted.

$\hat{q}_{\text{gender}}$	$\hat{q}_{\text{height}}$	$\hat{q}_{\text{bmi}}$	#
male	> 177.6	> 25	6
male	> 177.6	< 25	4
female	< 177.6	< 25	2
male	< 177.6	> 25	1
male	< 177.6	< 25	1
female	> 177.6	< 25	1

4.2 Daily Routine De-anonymization

In this section we present the results for re-identification of users based on their activity. Our best model achieves a 93.5% de-anonymization accuracy for the full 16-users D3 dataset. Again, we utilize time series of hourly (i) steps, (ii) distance, (ii) calories, and (ii) average hourly heart rate as features. Although the adversary is almost certain to find these basic parameters in a lifelogging dataset, there might be cases when she would like to utilize less training data. Therefore, we present the detection rates for the models that were not trained on the complete time series in Figure 5. There, for every number of time se-

ries features, ranging from 1 to 4, we chose the most relevant parameters in the following order: calories, average heart rate, distance, steps. It is worth noting that even with just time series of hourly calories the de-anonymization rate exceeds 80%, outperforming previous results (Marchioro et al., 2021). Furthermore, we report the re-identification results for lesser numbers of participants in D3 (Figure 6). We perform a Monte Carlo simulation, where for every number of users N, ranging from 2 to 15, we run 10 rounds of the experiment: we randomly select N participants from D3, and train the inference model. Then, we average the results for each value of N to get a final accuracy estimation. The extrapolation of our findings suggests that it may be possible to maintain high de-anonymization rate even for bigger datasets.

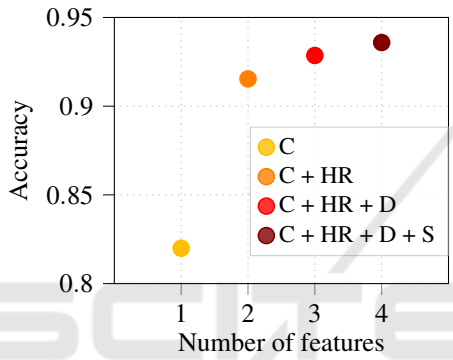


Figure 5: De-anonymization probability utilizing less time series features: C=calories, HR=heart rate, D=distance, S=steps.

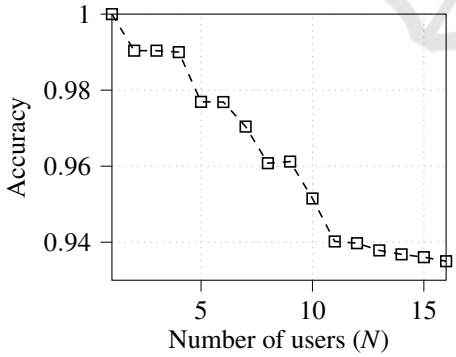


Figure 6: De-anonymization probability based on daily activity when subsampling users from the test dataset.

5 DISCUSSION

In this section we discuss the real-world applicability of our attacks, and possible ways to mitigate them.

Inferability of Personal Attributes. Previous sections suggest that daily activity data carry information about users' personal attributes such as gender, etc. Such information plays a fundamental role in the de-anonymization tasks, especially for groups with the imbalanced personal parameters distribution. An intuitive explanation for successful inference of such parameters is associated with the way Fitbit fitness attributes are produced: some (e.g., steps) are directly derived from sensor data, while others (e.g., distance, calories) are estimated from other personal information provided by the user[3]. Therefore, it is only natural that some of the data that we utilized as features might depend on physical parameters.

Are These Realistic Threat Models? A natural concern to rise would be the size of the datasets utilized for this work. However, we emphasize again that, in practice, the aggregated lifelogging datasets contain a very limited number of users as discussed in sections 1-2. Furthermore, grouping users with even limited amount of binary queries can considerably reduce the anonymity group for the targeted user. In some cases such reduction might reach a single person, meaning a 100% to de-anonymize the target. Even if the victim is not a minority individual, and cannot be detected very accurately with our threat model that is based on inferring physical characteristics, we have shown that they *still* can be successfully de-anonymized based on their daily routine, and activity patterns. In fact, the real-world adversary might choose a strategy of using the first threat model to target the minority users, and second otherwise. Naturally, it is likely that an attacker obtains the best re-identification results, combining both threat models. Given the flexibility of the possible attack flows, we believe it is possible to maintain high de-anonymization accuracy even for bigger datasets.

Possible Defense Mechanisms. Protecting participants in a public fitness dataset against our threat models might be way more complex than it appears. Applying traditional anonymization approaches, such as k-anonymity, to the quasi-identifiers would not be effective, since personal attributes are being inferred directly from the data. A natural solution would be to change the values of daily fitness time series, in order to confuse the inference models. For example, k-anonymity can be applied directly to the fitness information, but that would likely alter time series data beyond usability. Differential privacy (Dwork, 2006) is

[3]https://help.fitbit.com/articles/en_US/Help_article/1141.htm

another well-known solution to prevent membership inference in aggregated datasets. Differentially private mechanisms introduce noise in order to minimize changes in the data distribution caused by adding or removing a user. However, differential privacy typically cannot be applied to data collections as small as lifelogging datasets (Section 1).

More recent works have identified adversarial neural networks as a solution to protect time series data collected from smartphones (Malekzadeh et al., 2019). Such networks train a release mechanism that is used to "sanitize" the samples, concealing personal information. Their effectiveness on mobile sensor data suggests that they may also be used to anonymize fitness records from wearables.

6 CONCLUSION

We demonstrated that it is possible for the adversary to de-anonymize the records of anonymous users in an aggregated data collection, and uniquely re-identify minority individuals within the datasets based on their gender, height, and BMI.

We also showed that an adversary can de-anonymize all users (minority or majority) in the dataset based on their daily routine with 93.5% accuracy, if she has access to some of their fitness data.

Finally, we discussed how applying k-anonymity to quasi-identifiers (i.e., physical characteristics) would not guarantee users' privacy, since the adversary is still able to glean information on those attributes through the presented inference model.

REFERENCES

CCS Insight (2021). Healthy outlook for wearables as users focus on fitness and well-being.

Christovich, M. M. (2016). Why should we care what fitbit shares-a proposed statutroy solution to protect sensative personal fitness information. *Hastings Comm. & Ent. LJ*, 38:91.

Dong, Y., Hoover, A., Scisco, J., and Muth, E. (2012). A new method for measuring meal intake in humans via automated wrist motion tracking. *Applied psychophysiology and biofeedback*, 37(3):205–215.

Dwork, C. (2006). Differential privacy. In *International Colloquium on Automata, Languages, and Programming*, pages 1–12. Springer.

Furberg, R., Brinton, J., Keating, M., and Ortiz, A. (2016). Crowd-sourced Fitbit datasets 03.12.2016-05.12.2016.

Harris, J. A. and Benedict, F. G. (1918). A biometric study of human basal metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 4(12):370.

Hilts, A., Parsons, C., and Knockel, J. (2016). Every step you fake: A comparative analysis of fitness tracker privacy and security. *Open Effect Report*, 76(24):31–33.

Kelly, D., Curran, K., and Caulfield, B. (2017). Automatic prediction of health status using smartphone-derived behavior profiles. *IEEE journal of biomedical and health informatics*, 21(6):1750–1760.

Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):3–es.

Malekzadeh, M., Clegg, R. G., Cavallaro, A., and Haddadi, H. (2018). Protecting sensory data against sensitive inferences. In *Proceedings of the 1st Workshop on Privacy by Design in Distributed Systems*, pages 1–6.

Malekzadeh, M., Clegg, R. G., Cavallaro, A., and Haddadi, H. (2019). Mobile sensor data anonymization. In *Proceedings of the international conference on internet of things design and implementation*, pages 49–58.

Marchioro, T., Kazlouski, A., and Markatos, E. (2021). User identification from time series of fitness data. In *International Conference on Security and Cryptography (SECRYPT)*, pages 806–811.

OpenHumans (2016). Open humans fitbit connection. https://www.openhumans.org/activity/fitbit-connection.

Parate, A., Chiu, M.-C., Chadowitz, C., Ganesan, D., and Kalogerakis, E. (2014). Risq: Recognizing smoking gestures with inertial sensors on a wristband. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, pages 149–161.

Samarati, P. (2001). Protecting respondents identities in microdata release. *IEEE transactions on Knowledge and Data Engineering*, 13(6):1010–1027.

Sathyanarayana, A., Joty, S., Fernandez-Luque, L., Ofli, F., Srivastava, J., Elmagarmid, A., Arora, T., and Taheri, S. (2016). Sleep quality prediction from wearable data using deep learning. *JMIR mHealth and uHealth*, 4(4):e125.

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570.

Thambawita, V., Hicks, S., Borgli, H., Pettersen, S. A., Johansen, D., Johansen, H., Kupka, T., Stensland, H. K., Jha, D., Grønli, T.-M., and et al. (2020). Pmdata: A sports logging dataset.

Torre, I., Sanchez, O. R., Koceva, F., and Adorni, G. (2018). Supporting users to take informed decisions on privacy settings of personal devices. *Personal and Ubiquitous Computing*, 22(2):345–364.

Vitak, J., Liao, Y., Kumar, P., Zimmer, M., and Kritikos, K. (2018). Privacy attitudes and data valuation among fitness tracker users. In *International Conference on Information*, pages 229–239. Springer.

World, D. (2017, December 1). Height chart of men and women in different countries. disabled world. www.disabled-world.com/calculators-charts/height-chart.php. Online; Retrieved May 2, 2022.