Giridhari Venkatadri*, Elena Lucherini, Piotr Sapiezynski, and Alan Mislove

Investigating sources of PII used in Facebook's targeted advertising

Online social networking services have become the gateway to the Internet for millions of users, accumulating rich databases of user data that form the basis of their powerful advertising platforms. Today, these services frequently collect various kinds of personally identifying information (PII), such as phone numbers, email addresses, and names and dates of birth. Since this PII often represents extremely accurate, unique, and verified user data, these services have the incentive to exploit it for other purposes, including to provide advertisers with more accurate targeting. Indeed, most popular services have launched PII-based targeting features that allow advertisers to target users with ads directly by uploading the intended targets' PII. Unfortunately, these services often do not make such usage clear to users, and it is often impossible for users to determine how they are actually being targeted by advertisers.

In this paper, we focus on Facebook and investigate the sources of PII used for its PII-based targeted advertising feature. We develop a novel technique that uses Facebook's advertiser interface to check whether a given piece of PII can be used to target some Facebook user, and use this technique to study how Facebook's advertising service obtains users' PII. We investigate a range of potential sources of PII, finding that phone numbers and email addresses added as profile attributes, those provided for security purposes such as two-factor authentication, those provided to the Facebook Messenger app for the purpose of messaging, and those included in friends' uploaded contact databases are all used by Facebook to allow advertisers to target users. These findings hold despite all the relevant privacy controls on our test accounts being set to their most private settings. Overall, our paper highlights the need for the careful design of usable privacy controls for, and detailed disclosure about, the use of sensitive PII in targeted advertising.

***Corresponding Author: Giridhari Venkatadri:** Northeastern University, E-mail: venkatadri.g@husky.neu.edu
Elena Lucherini: Princeton University
Piotr Sapiezynski: Northeastern University
Alan Mislove: Northeastern University

1 Introduction

Users conduct an increasingly large fraction of their everyday activities online, often via online social network services such as Twitter and Facebook. By virtue of being free, these services have become extremely popular; this has allowed them to collect data about an extensive set of users. These services use this data for various purposes, most notably to build advertising platforms through which advertisers can target platform users.

In particular, these services collect significant amounts of *personally identifiable information* (PII)—information such as email addresses or phone numbers that uniquely identify users—for a variety of uses. For example, on Facebook, many of these uses are user-facing features: email addresses serve as login usernames, phone numbers allow users to find each other on Messenger, and users can "sync" their address books to find others they are not yet "friends" with.

However, there are other uses of PII that primarily benefit third parties. Most notably, many services have recently deployed PII-based targeting features to their advertising platforms [3, 37, 43], which allow advertisers to directly choose which users see their ads by providing a list of those users' PII. This feature—called *custom audiences* on Facebook's platform—is now popular among advertisers as it allows them to exploit the existing PII they have about their customers (such as email addresses, phone numbers, or names and addresses) and target them with advertisements.

Recent events have brought the issue of how user data is collected, used, and made available to third parties to the forefront. In particular, it was recently revealed that tens of millions of users' Facebook profile data was collected by an innocuous Facebook app, and then later shared with Cambridge Analytica (a data mining consultancy) for use in targeting political ads in the 2016 U.S. election [7]; custom audiences potentially played a significant role in accomplishing this targeting [18, 22, 25]. In response to the resulting uproar among both the press and lawmakers, Facebook changed certain aspects of how apps can collect data [1, 34].

While the Cambridge Analytica story received significant attention, the resulting privacy debate focused

largely on third-party apps and other such vectors of data leakage, and *not* on the advertising platform that these companies use to exploit such data. For example, even though Facebook removed the functionality that allowed *users* to find other users using phone numbers as part of the response to the Cambridge Analytica story [34], *advertisers* can still use phone numbers for targeting ads. Unfortunately, we have little understanding on how Facebook collects user PII, associates PII with user accounts, and makes PII available for use by advertisers via custom audiences.

In this paper, we address this situation by developing a novel methodology to study how Facebook obtains the PII that they use to provide custom audiences to advertisers. We test whether PII that Facebook obtains through a variety of methods (e.g., directly from the user, from two-factor authentication services, etc.) is used for targeted advertising, whether any such use is clearly disclosed to users, and whether controls are provided to users to help them limit such use.

Developing such a methodology presents two challenges: *First*, how do we verify whether PII added to an account has actually been used by Facebook for PII-based targeting? *Second*, how do we select "fresh" pieces of PII that are not already associated with some other Facebook user, in order to prevent incorrect inferences? We solve both of these challenges by developing a technique to check whether a given piece of PII can be used to target *some* Facebook user (i.e., is *targetable*).

Our technique exploits the size estimates that reveal how many users in a custom audience can be targeted with ads; these estimates are a fundamental feature of many advertising platforms, as they help advertisers budget their ad campaigns. We first reverse-engineer one such size estimate: *potential reach* [4], which reports the number of users in an audience who are active on a daily basis.[1] We show that Facebook obfuscates potential reach using a combination of rounding and noise seeded by the uploaded records.[2] Despite this obfuscation, we develop a technique of uploading lists of PII of consecutive sizes, adding "dummy" padding records to each list to get multiple samples at each size, and then using these samples to conclude whether the true number of matched users is different across consecutive sizes. We demonstrate that this approach is able to effectively negate the effect of Facebook's obfuscation, allowing us to check whether a single piece of PII can be used to target a Facebook user with ads via custom audiences.

We then use our technique to check *which* of a variety of potential sources of PII are actually used by Facebook to gather PII for targeted advertising. For example, if we wish to study whether phone numbers provided for two-factor authentication (2FA) are used for targeted advertising, we first obtain a new phone number. We then verify (using the technique above) that no user is currently targetable via this phone number; if so, we add it as a 2FA number to a control account. We then repeatedly check over the subsequent month (again using the technique above) to see whether the phone number becomes targetable. Finally, to verify our results if the number does become targetable, we run a controlled advertisement targeting the phone number and confirm that our ads are received by the control account.

We examine seven different sources of PII to see which are used for targeted advertising: (1) PII added directly to a user's Facebook profile, (2) PII provided to the Facebook Messenger app, (3) PII provided to WhatsApp, (4) PII shared with Facebook when sharing a phone's contacts, (5) PII uploaded by advertisers to target customers via custom audiences, (6) PII added to user accounts for 2FA, and (7) PII added for login alerts. We find that five of these result in the PII being used for advertising: all except for PII provided to WhatsApp and PII uploaded by advertisers.

Unfortunately, we find that Facebook does not directly disclose its PII practices beyond generic statements such as [13]:

> We use the information we have about you—including information about your interests, actions and connections—to select and personalize ads, offers and other sponsored content that we show you.

Worse, we found no privacy settings that directly let a user view or control which PII is used for advertising; indeed, we found that Facebook was using the above PII for advertising even if our control account user had set the existing PII-related privacy settings on to their most private configurations. Finally, some of these phone numbers that were usable to target users with did not even appear in Facebook's "Access Your

1 While these size estimates were reverse-engineered in prior work [40] and were found then to be computed by simple rounding, we found that Facebook now uses a more sophisticated way of obfuscating these size estimates (potentially to defend against the privacy vulnerabilities discovered by that work).

2 This was the case during the time period of our experiments in early 2018; as discussed in more detailed Section 5, Facebook has now temporarily removed these statistics, but has other related statistics that could likely be used.

Data" feature that allows users to download a copy of all of their Facebook data as a ZIP file.

Taken together, our results highlight the need to make the uses of PII collected clear to users, and to provide them with easy-to-use privacy controls over their PII. The sources of PII that we investigated, while being the most straightforward ones, are by no means exhaustive. However, since our method relies only on the provision of size estimates by PII-based advertising platforms, and since size estimates are an integral part of these platforms (as they are valuable to advertisers), methodology similar to ours can be potentially used to investigate other sources of PII and other services.

Ethics: All the experiments in this paper were conducted according to community ethical standards. All were performed only using the authors' accounts (or one fake account we created for this paper) and did not interact with other Facebook accounts in any way. When experimenting with the Facebook interface, we only used email addresses and phone numbers that we controlled, email addresses that Facebook already had, or publicly-available data. In no instance did we reveal any user PII that we did not already have, or disclose any PII to Facebook that Facebook did not already possess. We also ensured that we put a minimal load on Facebook's systems: we only created one fake Facebook advertiser account, made a limited number of API queries, and respected rate-limits.

Additionally, we responsibly disclosed our discovery of a method to check whether a given piece of PII is targetable to Facebook's security team. Facebook responded by stating this was not a security vulnerability [9] and closed our bug report:

> Enumeration vulnerabilities which demonstrate that a given e-mail address or mobile phone number is tied to "a Facebook account" are not eligible under the bug bounty program. This is true whether the endpoint returns only confirmed items, or both confirmed and unconfirmed. In absence of the user ID that the e-mail/mobile number is linked to, this behavior is considered extremely low risk.

Overall, we believe that any de minimis harm to Facebook as a result of our experiments is outweighed by the benefits to users in terms of increased transparency and understanding of how their PII is used.

2 Background

We begin by providing background on online social network advertising. In this paper, we focus on the Facebook advertising platform, as it is the largest and most mature. However, other competing services now offer similar features (e.g., PII-based user targeting), including Google's Customer Match [3] and Twitter's Tailored Audiences [37]. Thus, similar issues may exist on these sites as well; we leave a full exploration to future work.

2.1 PII-based targeting

Advertising on Facebook traditionally relied on selecting the *attributes* of the users to whom the advertiser wished to show ads. For example, an advertiser might specify that they wish to show ads to 30–35 year-old males who live in Chicago and who like a particular TV show. To allow the advertiser to place multiple ads with the same set of attributes, Facebook automatically creates a "group" of users after the advertiser selects the attributes of interest; these groups are called *audiences*.

Recently, Facebook—along with many other sites [40]—has introduced a new feature: *custom audiences*. Unlike traditional audiences, the advertiser is allowed to target *specific* users. To do so, the advertiser uploads user PII to Facebook; Facebook then matches the provided PII against platform users. Facebook currently allows users to upload 15 different types of PII, including phone numbers, email addresses, names, and even advertising tracking IDs [40]. Facebook then creates an audience consisting of the matched users and allows the advertiser to target this specific audience.

Figure 1 shows an example of the audience creation flow. In panel **A**, the advertiser is prompted to select the source of user information for targeting. The first option allows them to upload a list of users (e.g., existing customers).[3] In panel **B**, the advertiser is instructed on the types of data available for targeting, including email addresses, phone numbers, mobile advertising IDs, etc.

Audience statistics Facebook does not reveal to the advertiser *which* users were actually in the matched group, but it does provide two statistics: the total number of matched users, called the *audience size*; and the number of daily *active* matched users, called the *poten-*

3 The data is uploaded "in an encrypted format to maintain privacy" [12]; in reality, it hashed using SHA-256 with no salt.

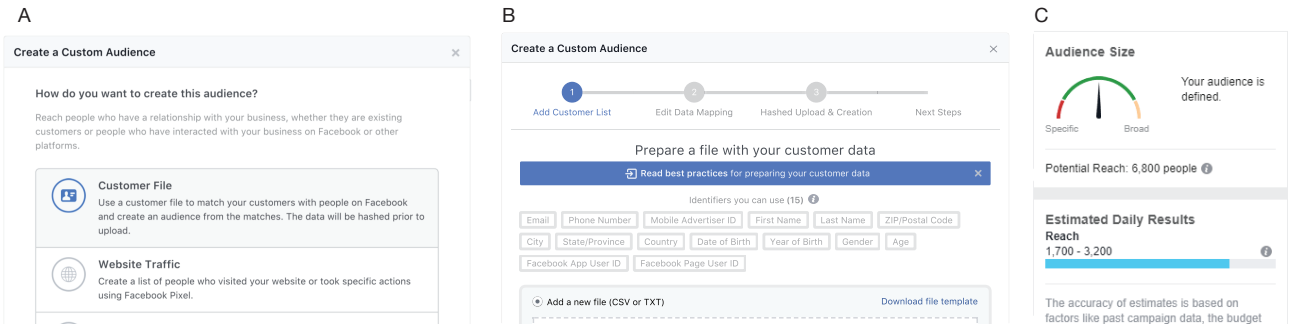

Fig. 1. Flow of creating custom audiences using the Facebook advertising site. The advertiser can select to upload data (panel **A**), and then can choose from among 15 supported types of PII (panel **B**). Once the data has been uploaded, the advertiser is provided with coarse-grained statistics about the users who matched (panel **C**).

tial reach [40]. An example of the reported potential reach is shown in Figure 1 panel **C**. Facebook also allows advertisers to *combine* multiple custom audiences, and will provide the potential reach of the combination; we refer to this as the *union* of the audiences.

We recently demonstrated that even these coarse-grained audience size approximations could allow advertisers to infer PII of particular Facebook users by observing changes in size statistics [40]. In brief, this attack worked because Facebook "deduplicated" PII that referred to the same underlying Facebook user in a list of uploaded PII. Facebook claims to have mitigated this issue by refusing to provide audience size and potential reach statistics when the advertiser uploads PII of multiple types; irrespective of whether this is the case, we show in the next section that we are still able to infer whether specific uploaded PII is targetable.

2.2 Site features

There are a few features that we investigate as potential sources of PII for advertising.

Profiles Users are allowed to provide PII such as email addresses and phone numbers as part of their basic profile, both to serve as their login username and to be revealed to friends. Such user-reported PII could be used to match against advertiser-uploaded PII.

Login Alerts If the user opts into Login Alerts, they are notified whenever anyone successfully logs in to the account from a "new" device. The alerts can be presented as a Facebook notification, a Messenger notification, an email, or an SMS. The two latter channels require the user to provide the email address or phone number (i.e., PII) to which the notification should be sent; as we will see later, this information could then be used to match against advertiser-uploaded PII.

Two-Factor Authentication (2FA) Facebook allows adding a variety of second security factor (a *what-you-have* factor) authentication methods: SMS messages, USB security keys, code generators, and one-time-use recovery codes. The most commonly used of these is the SMS message, which requires a user to provide Facebook with a phone number to send the SMS to. Similar to login alerts, this PII could then be used to match against advertiser-uploaded PII.

Address book synchronization Facebook users can find their friends on the platform by allowing the Facebook app to access their phone's address book. Each contact in the address book can have multiple pieces of PII, for example the name, email address, and phone number. Hence, Facebook could potentially match contacts based on partial PII (just the email address), but still learn new PII (a phone number of a person whose email address is known to the platform).

WhatsApp Users are identified in WhatsApp by their phone numbers. If a user has both WhatsApp and Facebook apps installed on the same (Android) phone, Facebook could use the Android advertising ID to learn that the two disconnected accounts belong to the same user, and thus associate the phone number with the Facebook account as well.

Messenger Upon installation of the Messenger app, the user is prompted to upload their address book (potentially leaking contacts' PII) and to use it as the default SMS app. Granting the latter permission reveals the user's phone number to Facebook.

3 Methodology

We now develop a methodology to check whether Facebook uses PII from a given source for targeted advertising.

3.1 Datasets and setup

In order to reverse-engineer the size estimates provided by Facebook's PII-based advertising platform, similar to prior work [40], we collect 103 emails and phone numbers corresponding to friends and family who have Facebook accounts and had previously provided their PII to Facebook (i.e., they had already done so; we did not ask them to upload it). There were no requirements (activity or otherwise) asked of these users. Thus, these users were not affected in any way by our experiments.

To differentiate between PII which matches Facebook users and PII which does not, we also create *dummy PII* designed not to match any Facebook user. We generate dummy phone numbers by adding a random sequence of 20 digits to the Italian country code (+39). Since Italian phone numbers do not exceed 12 digits, these dummy numbers cannot correspond to any Facebook user. Similarly, we generate dummy email addresses by using randomly generated strings of alphabets as usernames and (long) dummy domain names.[4] We then automate the process of uploading lists of PII to create custom audiences, and of collecting potential reach estimates, using scripts that make appropriate calls to Facebook's marketing API [21].

3.2 Reverse-engineering potential reach

As described in Section 2, Facebook's advertising platform provides two audience size estimates; for this paper, we only use the potential reach, which measures the number of active users in the audience. Below, we reverse-engineer how the displayed potential reach estimates are computed.

3.2.1 Are size estimates obfuscated?

Our prior work [40] demonstrated that Facebook's advertising platform obfuscated the potential reach by simple rounding; to do so, we showed that the potential reach estimates were granular (rounded in steps of 10 up to 1,000),[5] consistent over short periods of time while occasionally varying over longer time periods (which is expected since some active users might become inactive and vice versa), and monotonic. However, Facebook's interface changed significantly since that work was conducted; thus, we revisit these findings to see if Facebook still obfuscates size estimates via simple rounding.

Granularity: We create 10,000 different custom audiences by uploading sets of varying sizes containing either phone numbers or email addresses (from the 103 that we collected), and obtain their corresponding potential reach estimates. Consistent with prior work [40], we find that the estimates are still granular and increase in steps of 10, always returning one of $\{20, 30, 40, ..., 80\}$.[6]

Consistency: To check consistency of potential reach estimates over short periods of time, we create two audiences by uploading 70 and 89 phone numbers respectively; for each audience, we make 1,000 potential reach queries back-to-back. We found that all the 1,000 queries returned the same potential reach for each audience, with their estimates being 40 and 80 respectively. We repeated the above experiment for various lists, both of phone numbers and of emails, and at different times, and found that the potential reach estimates for a given audience were always consistent over short periods of time. This is also consistent with prior work [40].

To check the consistency of potential reach estimates over longer periods of time, we take three audiences created by uploading 60, 80, and 103 phone numbers respectively, and repeatedly obtain the potential reach for each audience every five minutes, over a period of around 14 hours, giving us 164 samples in total for each audience. We find that the estimates were consistent over this period of time. with values of 40, 60, and 80 respectively. We repeated the above experiment for other audiences and across longer periods of time

4 To confirm that the dummy PII we created do not correspond to any Facebook user, we created two audiences containing 1,000 dummy phone numbers and 1,000 dummy email addresses; both audiences had a potential reach of 20 (the smallest value it can take), meaning that the dummy PII indeed do not correspond to any Facebook user.

5 We do not discuss or investigate larger audience sizes as they are not necessary for our paper.

6 It is important to note that the potential reach may not always be 100 even though we uploaded 100 PII records, both because we uploaded subsets of varying size and because potential reach only counts "daily active" users.

and found that the size estimates were generally consistent, sometimes changing over a period of hours. This is also consistent with prior findings, and is expected as whether a given user is "daily active" (and counts towards potential reach) may change over time.

Finally, to check the consistency of potential reach across multiple uploads of the same list of PII, we repeatedly upload a list of 70 phone numbers 100 times over a period of three hours, and obtain the corresponding potential reach estimates; we find that the estimates are generally consistent across uploads, with 99 of the custom audiences having a potential reach of 40, with only one having a different potential reach of 50.[7]

Monotonicity: Prior work [40] found that the potential reach was monotonic, meaning adding additional records to an uploaded list would never reduce the potential reach. To check whether the potential reach estimates are still monotonic, we upload a series of lists of phone numbers, starting at 70 numbers and successively adding one number until we reach 89 numbers. Surprisingly, we find that the potential reach *does not* increase monotonically! For example, uploading a list of 77 phone numbers resulted in a potential reach of 70; adding three more records to these 77 and uploading the resulting list resulted in a potential reach of 50. This indicates Facebook's potential reach computation has changed, and that they are likely obfuscating the potential reach estimates by randomly perturbing them. We repeated this experiment with other series of lists of phone numbers and email addresses, and found that similar lack of monotonicity holds.

Summary: We find that the potential reach estimates remain granular, rounded to the nearest 10 (for the range of values that we observed), and remain consistent for a given audience across short periods of time, as observed in prior work [40]. However, we find that the potential reach estimates are no longer monotonic, indicating that Facebook might be additionally perturbing potential reach estimates by randomly perturbing them with noise. Therefore, we move on to reverse-engineer the updated—potentially more sophisticated—way in which Facebook obfuscates potential reach estimates.

7 While we are not sure about why this one upload resulted in a different value, we believe this could either be because of an occasional error in the process of creating an audience, or because of the variation of potential reach over longer periods of time.

3.2.2 Properties of noisy estimates

Since Facebook appears to be using noise to perturb the potential reach estimates, we move on to study how the noise is seeded, and to characterize the relationship of the noisy estimates corresponding to a given custom audience with the true value.

What seeds the noise? Since the potential reach estimates are consistent across multiple repeated queries to the same custom audience, this indicates that a fresh sample of noise is not generated corresponding to each query, and that the noise is fixed for a given custom audience (perhaps to limit a malicious advertiser's ability to generate multiple noise samples). Additionally, since multiple uploads of the same list of PII records have the same potential reach, this indicates that the same seed is used to compute the noise sample whenever a given list of PII records is uploaded (indicating that this seed is computed using the list of PII records uploaded, for example by using a hash of the list contents).

In order to check whether all the PII records in a given list are used to determine this seed, or whether only records that match some Facebook user are used, we take a list of 60 phone numbers and upload it 400 times, each time with a different dummy phone number added (i.e., a phone number that we know cannot match any user). This gives us 400 custom audiences, each with the same set of users (since they were created using the same list of valid PII records), which were created with different lists of PII records (since each list contains a different dummy record). We find that the potential reach varies across the audiences, with values 20 (appearing once), 30 (appearing 42 times), 40 (appearing 192 times) and 50 (appearing 165 times). We find that the result holds even if we separately create one audience corresponding to the 60 phone numbers, create 400 audiences corresponding to one different dummy record each, and then dynamically ask for the potential reach of the *union* of the large audience with each of the dummy audiences. This result indicates that Facebook is considering *all* the PII records uploaded when deterministically calculating the noise to add, regardless of whether they are valid records or not.

Summary: We find that Facebook obfuscates potential reach estimates corresponding to a given custom audience using a fixed noise value; the seed for this noise is computed based on the list of (both valid and invalid) PII records uploaded. However, this suggests a method to obtain multiple noisy estimates corresponding to a given audience, and potentially overcome the effect of

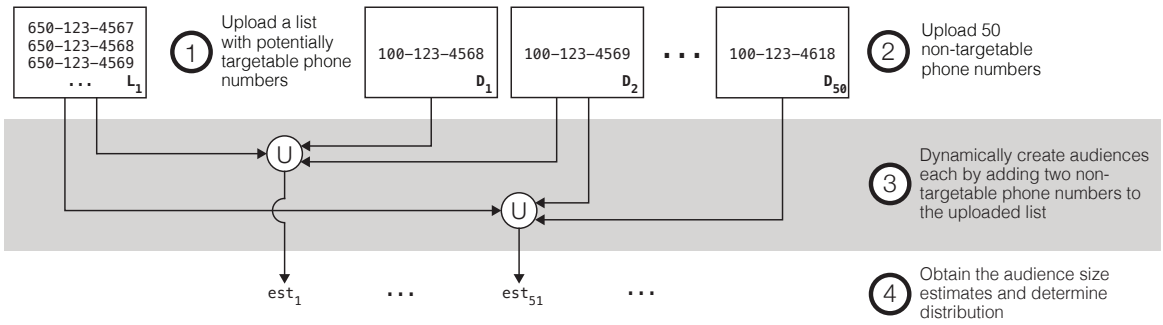

Fig. 2. Process of using combinations of a small number of audiences to obtain a large number of samples. In our case, we combined the target audience with pairs of a set of 50 audiences, resulting in 1,225 samples for the target.

noise: upload the same list of PII records multiple times with a different dummy record added each time and obtain the corresponding potential reach values. We can then examine the *distribution* of potential reach values to say something about the true underlying value.

Obtaining a large number of samples To measure the distribution of potential reach values, we need a way to easily obtain a large number of samples of noisy estimates without having to upload a large number of dummy audiences (since Facebook only allows us to maintain 500 custom audiences in a given advertising account). To accomplish this, we extend the idea of combining dummy audiences proposed in the previous section by creating 50 audiences with a different dummy phone number each, and then dynamically taking the union of *two dummy audiences* at a time with the given custom audience, as illustrated in Figure 2.

Since the list of PII records corresponding to a given combination of audiences is different (since each combination corresponds to a different combination of dummy records), each combination of two dummy audiences should give us a different sample of the noisy estimate. Using 50 dummy audiences gives us 1,225 samples corresponding to all possible combinations of dummy audience pairs; it takes up to 20 minutes to obtain all samples once all the audiences are uploaded and ready (at the rate of about a query a second).

Distribution of noisy estimates To characterize the distribution of noisy estimates, we upload consecutive lists of 70 to 90 phone numbers, each with one phone number beyond the previous one; we then obtain the distribution of potential reach for each of the audiences (1,225 samples per audience), and plot the variation in the histogram of the distribution against the number of phone numbers used in the first part of Figure 3. We observe a number of interesting trends:

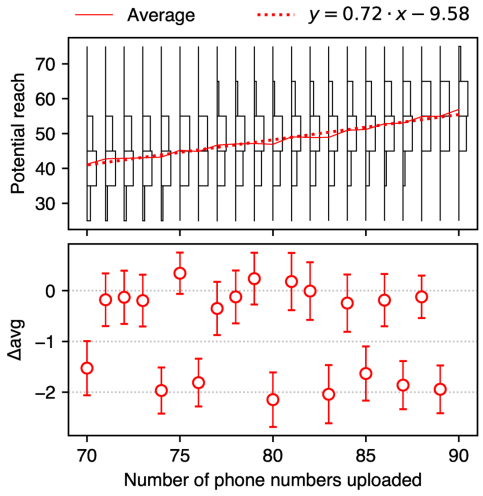

Fig. 3. Figure showing how the distribution of noisy potential reach values varies with with uploaded sets of phone numbers of increasing size. The first part of the figure shows how the histogram and average of potential reach values varies, while the second part of the figure shows how the average changes between subsequent distributions. Error bars correspond to confidence intervals at 95% confidence.

Is noise bounded? We notice from the first part of Figure 3 that the observed potential reach values corresponding to a given set of phone numbers is always drawn from a set of three consecutive multiples of 10, or two consecutive multiples of 10; for example, the distribution corresponding to the set of 70 uploaded phone numbers has values in the set $\{30, 40, 50\}$.[8] We experimented with phone number lists of various sizes (up

8 For four of the phone number lists, we occasionally notice a very small number of samples with an outlier value (the maximum number of samples observed being 4). We believe this might be because of occasional inconsistencies when combining a custom audience with different dummy records.

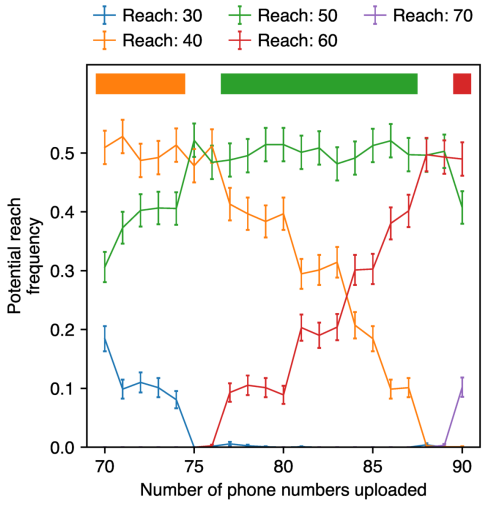

Fig. 4. Figure showing how the frequency of observed potential reach values changes with uploaded sets of phone numbers of increasing size. Whenever there are three values observed for a given set of phone numbers, the bar at the top shows the median value. Error bars correspond to confidence intervals at 95% confidence.

to 100) and observed that this result held irrespective of the size of the phone number list. This experiment shows that the noise is bounded, that noisy estimates are drawn from a range of thirty consecutive values or lesser (since each potential reach value could potentially correspond to one of ten unrounded values), and that these bounds do not depend on the magnitude of the actual potential reach (at least within the range of values that we study).

Is noise added before or after rounding? Having observed both that the true value is obfuscated via rounding (in steps of ten) and perturbed by noise, we move on to examine whether the true value is first perturbed with noise and then rounded, or whether it is first rounded and then perturbed with noise. To do so, we study how the distribution of observed values shifts as the size of the corresponding list of phone numbers increases. In the upper panel of Figure 3 we see how the histogram of potential reach values shifts towards higher values as the size of the phone number list increases; we see that the frequency with which any potential reach value occurs changes in discrete steps as the size of the phone number list increases.

To further characterize the steps by which these frequencies change, Figure 4 shows how the frequency of occurrence of a particular potential reach value in a distribution varies with the size of the corresponding list of phone numbers. From the figure, we see that frequencies of occurrence change in steps of uniform size (of about 0.1). *First*, note that such uniform steps are what would be expected if noise is added to (or subtracted from) the (rounded or unrounded) true value, as opposed to say multiplied, in which case the steps would be non-uniform. *Second*, if the noise is added to the true value after rounding it (in steps of ten), then the distribution would shift if and only if the underlying rounded value had changed (by ten); thus, we would expect that with every "shift" in the distribution the set of observed values would shift by 10. For example, assume that the value after rounding (before adding noise) is 60 and the corresponding set of observed values is $\{30, 40, 50\}$. This distribution would only "shift" when the value after rounding "shifts" to the next multiple of 10 (to 70 say), in which case the corresponding set of observed values would be expected to be $\{40, 50, 60\}$. However, this is contrary to the much finer steps with which the observed distribution shifts, showing that noise is added *before* rounding.

Is distribution uniform or non-uniform? To study whether the noise added is uniformly drawn from uniformly spaced values, we study how the frequency of occurrences of the median of the three observed potential reach values changes as the true value increases (when there are only two observed potential values, either could be considered the median). For example, assume the noise added is uniformly drawn from a range of m contiguous values, such as from the set $\{0, \cdots, m-1\}$ (where $m \leq 30$). We would expect that of the range of contiguous values obtained by adding the noise to the true value, the smallest few values will be rounded to the smallest of the three observed potential reach values, that the next ten values would be rounded to the median of the three observed potential reach values, and the remaining (largest) values would be rounded to the largest of the three observed potential reach values. Whatever the true value, we would therefore expect that the median observed potential reach would always have exactly ten distinct noise values corresponding to it. Therefore, assuming the noise added is uniformly distributed over $\{0, \cdots, m-1\}$, the expected frequency of occurrences of the median potential reach is $\frac{10}{m}$ irrespective of the true value. On the other hand, if the noise added is non-uniformly distributed, the expected frequency of occurrences of the median potential reach would change with the true value. Similarly, if the noise is drawn from values that are non-uniformly spaced, the expected frequency of occurrences of the median potential reach would change with the true value.

To study this, we show the median potential reach value for each size of phone number list in Figure 4 by indicating the color of the appropriate line (we only show this when there are three distinct median values observed).[9] We observe that the frequency of the median potential reach remains constant (around 0.5) regardless of the increase in number of phone numbers uploaded, and despite the fact that the frequency of the other two values of potential reach show multiple changes over the range shown; this shows that the distribution of noise values is uniformly distributed (and that the values are uniformly spaced). Besides, since the expected frequency of occurrences of the median potential reach is $\frac{10}{m}$, we can determine that Facebook has chosen $m = 20$ and therefore that the noise is uniformly distributed between 0 and 20.

However, if the noise was indeed uniformly distributed over twenty consecutive values, whenever the true value increases by one, we would expect the frequency of the smallest observed potential reach value to decrease in steps of $\frac{1}{20}$ (i.e., of 0.05). However, as previously observed, the step sizes observed in figure 4 are close to 0.1, approximately double the expected value.

Investigating the unexpected step size: To investigate why the frequencies of observed potential reach values change in steps twice as large as expected (i.e, in steps of 0.1 rather than 0.05), we check whether the true value of the potential reach (obtained by averaging out the different samples) increases in steps of one as we increase the number of phone numbers uploaded. The lower panel of Figure 3 shows the changes in the average of the observed potential reach values between consecutive sizes of phone number lists. We see that all non-zero changes are close to two in magnitude, showing that the true value increases in steps of two, rather than one as expected, potentially showing that the true value is first rounded in steps of two (before adding noise and rounding in steps of ten).

To further confirm this, we upload a list of 61 phone numbers, and similarly create six custom audiences containing six different phone numbers respectively (one each) corresponding to users we know to be active Facebook users. We obtain the distribution of potential reach estimates corresponding to the 61 phone numbers; we then take a union of the audience with the audiences corresponding to each of the six phone numbers, adding in each one by one and finding the distribution of potential reach. We find that the distribution shifted with the addition of the first phone number, did not shift with the addition of the second phone number, and so on, shifting only with every alternate phone number added. Repeating the experiment with different phone numbers, we also find that distribution shifted with the addition of the first phone number, irrespective of which of the six phone numbers were chosen as the first phone number. This confirms that Facebook rounds the true value of the potential reach in steps of two, before obfuscating it further.

Summary Taken together, we find that Facebook is first rounding the true values of the potential reach estimates in steps of two, then adding (or subtracting) uniform pseudorandom noise seeded by the uploaded PII records and drawn from a range of 20 consecutive values, and finally rounding the result in steps of ten. Given this understanding of how potential reach is calculated, we can now revisit our original goal of determining whether a piece of uploaded PII can be used to target a Facebook user (i.e., is targetable).

3.3 Determining whether PII is targetable

Since Facebook first rounds the true value of the potential reach prior to adding noise (and then subsequently rounding the resulting value in steps of ten), we need to overcome the first layer of rounding by finding a *threshold* audience A_t whose true size falls right on the rounding threshold (i.e., adding another user to which would cause the value to be rounded to the next higher value). This idea of finding a threshold audience is adapted from our prior work [40].

Finding a threshold audience: To find a threshold audience, we first upload a series of PII lists to Facebook (call them $\{L_1, L_2, ..., L_n\}$), where each list consists of the previous list with *one* record added to it. We then check the potential reach distributions for the resulting audiences $\{A_1, A_2, ..., A_n\}$, and find an audience A_t such that the distributions for A_t and A_{t+1} are different. A_t is then our threshold audience (if A_t was not a threshold audience, the true size estimates of A_t and A_{t+1} would have been rounded to the same value, leading to identical distributions for the potential reach). In all our experiments in the previous section,

9 Figure 4 also sometimes shows a fourth observed potential reach value for some sizes of phone number lists. As previously described, we believe these to be because of occasional inconsistencies when combining different dummy records with a given custom audience; we disregard these when finding the median bin.

we noticed that the change in the number of occurrences (out of 1,225 samples) of the lowest observed potential reach estimate across consecutive PII lists was either very small (with no variation of more than 60), or large (never smaller than 90). Therefore, to check whether the distribution shifts between A_t and A_{t+1}, we check whether the number of occurrences of the smallest observed potential reach estimate drops by more than 90 (in expectation, we would expect a shift to cause a drop of 123 in the lowest bucket, or 10% of the 1,225 samples).

Checking whether PII is targetable: In order to check whether a given PII V is targetable, we compare the potential reach distributions of A_t versus $A_t \bigcup V$. If these come from different underlying distributions, then V matches an active Facebook user and is targetable (as adding V changed the distribution), else not. We check whether the distribution shifts between A_t and $A_t \bigcup V$ in a similar manner as above, checking whether the number of occurrences of the smallest observed potential reach estimate drops by more than 90.

Validation: To validate the above methodology, we generate ten dummy phone numbers and email addresses and check whether they can be used to target some user. We then check for three phones and two email addresses belonging to the authors (with active Facebook accounts) whether they can be used to target some user on Facebook. Using the technique proposed in this section, we find that none of the dummy records are targetable, while all of the PII corresponding to the authors are targetable.

3.4 Determining if a source of PII is used

We now describe a methodology that uses the technique developed in the previous section to check whether PII gathered by Facebook via a given source is actually used by Facebook for PII-based advertising. The methodology can be summarized as follows:

1. Pick a PII (e.g., a new phone number) that we control (call it the *test PII*) to use for the experiment. Check whether the test PII is targetable to begin with; if so, then it is already associated with some Facebook account (and thus might interfere with the experiment); pick another PII instead.
2. Take a Facebook account that we control (call it the *control account*) and the test PII from the previous step. Using the given source from which Facebook gathers PII about users, provide the test PII in a way that allows Facebook to associate it to the control account. This could be direct (e.g., adding the test PII directly to the control Facebook account as a 2FA number); or indirect (e.g. syncing a contact list containing the test PII from some other Facebook account so that Facebook can link the test PII to the control account). We describe in detail how we do this for different sources in Section 4.3.
3. Check daily over the next month whether the test PII becomes targetable. If it does so, confirm that it is associated with the control account by running an ad targeting the test PII; verify that the control account receives it. [10] If so, we can conclude that the given source is a vector for PII-based advertising.

4 Experimental Results

We continue by using the methodology developed above to investigate which of various potential sources of PII are actually used by Facebook to gather information for their PII-based advertising feature.

4.1 Facebook's data use policy

We first analyze Facebook's data use policy [13] (last revised on April 19, 2018) to understand what it reveals to users about the potential uses of PII, and about the sources from which PII is collected. Facebook's data policy covers the information it processes to support what it terms the "Facebook Products" or "Products" [42], which include services offered to users such as Facebook (including the mobile app), Messenger, Instagram etc.; and services offered to business partners, such as the Facebook Business Tools [38] (including various APIs, Facebook social plugins, etc.)

Other companies owned by Facebook [39] (currently eight are listed), such as WhatsApp, have their own privacy policies. Here, we focus on the information that is disclosed in Facebook's data policy.

Potential uses of PII First, we note that Facebook's data policy [13] describes the potential uses of PII collected only at a high level, and in general does not differentiate among different types of information or

10 This may not be guaranteed to succeed owing to the complexity of the ad delivery process [23]

the sources from which it is obtained. Regarding advertising, it does not directly refer to PII, but says:

> ... we use the information we have about you—including information about your interests, actions and connections—to select and personalize ads, offers and other sponsored content that we show you.

Potential sources of PII To understand potential sources of PII that Facebook could use, we analyze the sources of information listed in Facebook's data policy and describe how various sources listed there could be potentially used to collect PII. In general, we find that the sources of information are also described only at a high level, making it hard for users to understand which of them could be potentially used to collect PII, and what PII might be collected from each source.

"Information and content you provide" The policy mentions that Facebook collects "the content, communications and other information you provide when you use our Products, including when you sign up for an account, create or share content, and message or communicate with others." This indicates that PII directly provided to Facebook (e.g., the email address or phone number you use when you sign up to an account), or PII mentioned in messages with other users might potentially be collected and used for advertising.

"Things others do and information they provide" The policy mentions that Facebook collects "information that other people provide when they use our Products. This can include information about you", such as when they "upload, sync or import your contact information."; besides, the policy also mentions that contact information collected from such uploading, syncing, or importing, is used for any of the purposes listed in the policy (of which advertising is one). This indicates that PII provided about you to Facebook by other users might potentially be collected by Facebook and used for advertising; in our context, this is particularly worrying as the user may not even be aware that such PII about them has been collected by Facebook and is being used to target advertisements to them.

"Device information" The policy mentions that Facebook collects device information, including PII such as location and device IDs; and connection information such as language and mobile phone number.

"Information from partners" The policy mentions that Facebook receives information "about your activities off Facebook" and "about your online and offline actions and purchases" from third-party partners; while the policy mentions that PII is never shared with advertising partners (without user permission), or with measurement or analytics partners, it never mentions whether PII is received from these partners. Thus, PII about users could potentially also be obtained by Facebook from third-party partners (such as advertisers, app developers, third-party data providers etc.)

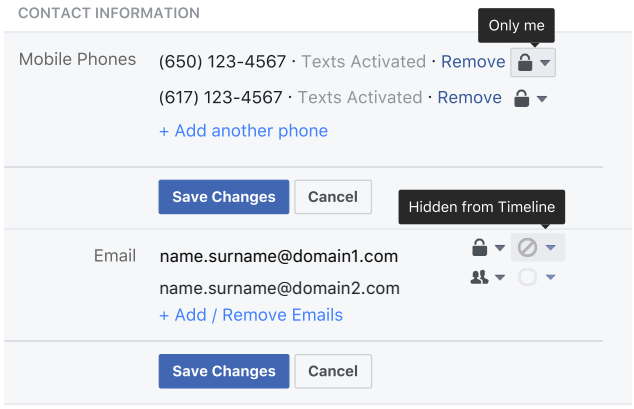

Fig. 5. Editing PII using the Facebook interface. The user can decide who to make the information available to and whether or not it should appear on their timeline.

Summary Facebook's data use policy reveals that a variety of potential sources of information, including sources where the user is not directly involved, could be used by Facebook to collect PII for advertising. However, we find that the data use policy describes the sources of information at a high level, making it hard for a user to understand which sources might be used to collect PII. Moreover, the policy simply mentions that all collected information might be used to target advertisements; this is likely insufficient for users to understand what sources of PII are used for targeted advertising.

4.2 Privacy controls for PII

We examined Facebook's interface and found that only the three following privacy options help control the usage of PII (we limit ourselves to high-fidelity PII: email addresses and phone numbers). *First*, users can specify who can see each PII listed on their profiles, the current list of possible general settings being: Public, Friends, Only Me; see Figure 5. In addition, users can specify a custom list of users who can see or not see the PII, or choose from preset groups of people (computed by Facebook) who match the user's workplace, location, university, etc. We call this the *profile privacy control*.

Second, Facebook allows users to restrict the set of users who can search for them using their email address

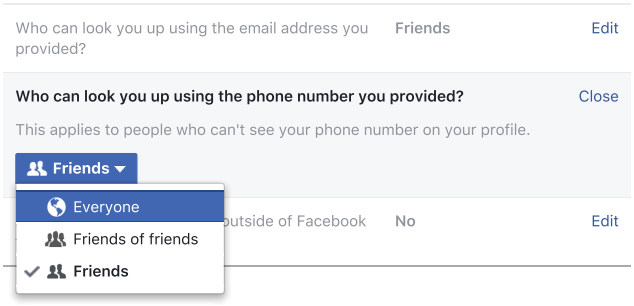

Fig. 6. In the privacy settings the user can decide who can look their profile up using the provided email address and phone number. The most restrictive option available is "Friends". We find that even when the user sets the PII visibility to "Only me" and searchability to "Friends", the advertisers can still use that bit of information for targeting.

or phone numbers; users can choose from the following options: Everyone, Friends of Friends, and Friends, see Figure 6.[11] We call this the *lookup privacy control*. Note that this control does not refer to any particular phone number or email address; it is one global setting for phone numbers and one for email addresses.

Third, on the ads preferences page [11], Facebook shows users a list of advertisers who have included them in a custom audience using their contact information. Users can opt out of receiving ads from individual advertisers listed here; however, they cannot see what PII is used by each advertiser. Additionally, Facebook does not let users directly control which PII is used to target advertisements to them.

4.3 PII sources for PII-based advertising

We move on to use the methodology proposed in Section 3 to study which of a number of potential sources of PII are actually used in PII-based advertising.

4.3.1 Setup

In order to obtain phone numbers to use for our experiments, we purchased SIM cards and plans from various mobile operators. We verified that some of the numbers were already targetable (as per our methodology proposed in Section 3); we discarded those and used only the numbers that were not targetable before our experiments. In addition, we used other email addresses belonging to the authors which they had not previously provided to Facebook (and which were similarly double-checked to not be associated with active Facebook accounts). We use the accounts of the three authors with active Facebook accounts for all our experiments. We performed a factory reset on the Android phone we used for these experiments before inserting each new SIM card, in order to wipe out any context that might lead to interference between experiments.[12]

4.3.2 PII provided as profile data

Facebook allows users to add contact information (email addresses and phone numbers) on their profiles. While any arbitrary email address or phone number can be added, it is not displayed to other users unless verified (through a confirmation email or confirmation SMS message, respectively). Since this is the most direct and explicit way of providing PII, we first study this to obtain a baseline estimate of how quickly Facebook makes newly collected PII available for targeted advertising.

We added and verified an email address and phone number to one author's account, and find that they both became targetable within six days. We also added an unverified email address and a phone number to one of the author's accounts (i.e., we did not complete the SMS/email verification process), and found that neither the email address nor the phone number became targetable after one month, suggesting that only verified phone numbers or email addresses are used for advertising.

Note that, for the purposes of this baseline experiment, we had set the least restrictive options of both the profile privacy control and the lookup privacy control. For all remaining experiments, we assume that a user is privacy-conscious and turns both the PII-level privacy controls to their most restrictive settings.

Disclosure and privacy controls: When users add mobile phone numbers directly to their profile, no information about potential uses of that number is directly disclosed to them, as shown by Figure 7 (panels A and C show the interfaces for adding phone numbers on the

11 Recently, Facebook changed its policy to disable the ability to look users via email addresses or phone numbers in response to data leakage attacks. However, these controls still remain.

12 While Facebook could potentially use the device's immutable identifiers, such as the IMEI number, to link data obtained across various factory resets, this is unlikely as it contravenes Google's Android developer best practices [6].

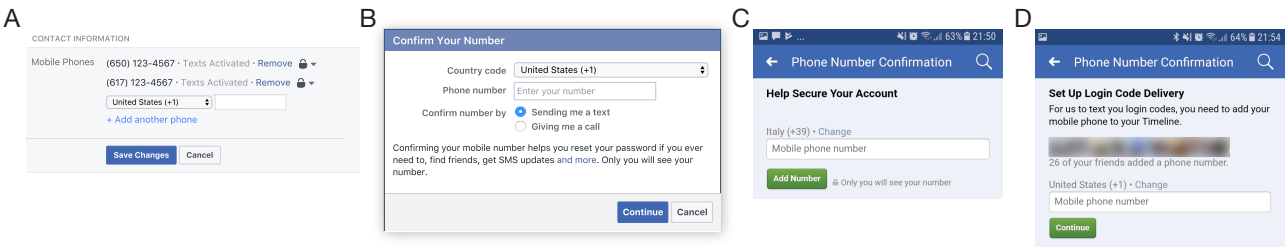

Fig. 7. Screenshots of interfaces used by users to add mobile phone numbers on Facebook's main website (A and B) and on Facebook's mobile app (C and D). While interfaces A and C come up when a user directly adds phone numbers to his Facebook profile, interfaces B and D arise when Facebook adds a phone number for a security feature.

website and the Facebook app respectively); the same holds for email addresses. Thus, users adding contact information for their friends' convenience may not be aware that their PII will then be used for targeting ads.

4.3.3 PII provided for security

We move on to examine whether PII provided by users for security purposes such as two-factor authentication (2FA) or login alerts are used for targeted advertising. Users may naturally provide this data with only security purposes in mind; if used for advertising, this may significantly violate a user's privacy expectations.

Two-factor authentication We added and verified a phone number for 2FA to one of the authors' accounts. We found that the phone number became targetable after 22 days, showing that a phone number provided for 2FA was indeed used for PII-based advertising, despite our account having set the privacy controls to the most restrictive choices.

Unrecognized login alerts Facebook allows users to add email addresses or phone numbers to receive alerts about logins from unrecognized devices. We added a phone number and an email address to an author's account to receive login alerts, and found that both the email address and phone number became targetable after 17 days.

Disclosure and privacy controls: Information about potential uses of a mobile phone number added for security purposes is only disclosed to users when adding a number from the Facebook website (and not from the Facebook mobile app). This can be seen from panels B and D of Figure 7, which show the interface for adding mobile phone numbers for security features using the website or app, respectively (no disclosure about potential uses happens elsewhere during the process). The interface informs users that "confirming your mobile number helps you reset your password if you ever need to, find friends, get SMS updates and more. Only you will see your number." The text "and more" in the above is hyperlinked to the Facebook's data policies page, as discussed in Section 4.1. There is no disclosure on either the website or the mobile app when email addresses are added to receive unrecognized login alerts. Finally, as with adding PII to the profile, there is no indication to the users that there exist other relevant privacy controls that they might want to revisit. Thus, it is highly likely that users are unaware that enabling security features through this interface will enable advertisers to target them directly with ads.

4.3.4 PII provided to other Facebook services

We move on to study whether PII provided on services owned by Facebook other than the main website and app is used for advertising.

Facebook Messenger Users must provide and verify a mobile phone number to the Facebook Messenger app if they want to use its SMS functionalities. We installed the Facebook Messenger app on a freshly wiped phone, added a phone number to it (verified with an SMS message), and checked whether the phone number became targetable. We find that the phone number did indeed become targetable after nine days. Again, use of these phone numbers for targeted advertising can potentially be counter-intuitive to users and violate their privacy expectations, since the phone number is provided with a specific purpose in mind (SMS messaging), and in the specific context of the Facebook Messenger app.

Disclosure: The first page of the process of adding a phone to the Messenger app discloses to the user that setting up Messenger "lets Friends find each other on Facebook and helps us create a better experience for everyone." However, apart from this generic description,

no other details are provided to the user about potential uses of the data collected.

WhatsApp Users are generally identified by their phone numbers on WhatsApp. To study whether these numbers are used for advertising, we first installed the Facebook app on a freshly wiped phone and logged in with one of the authors' accounts. We then installed the WhatsApp app on the same phone, providing a new phone number. We found that the new phone number did *not* become targetable even after a month. Note that our experimental setup is not exhaustive, and there may be other situations when Facebook would use WhatsApp phone numbers that we did not consider.

Disclosure: The first page of the process of adding a phone number to the WhatsApp app includes a link to WhatsApp's Terms of Service and Privacy Policy; these make a generic statement that Facebook "may use information from us to improve your experiences within their services" including for showing ads.

4.3.5 PII obtained without user's knowledge

Finally, we investigate whether PII obtained without a user's knowledge, such as by some other user syncing their phone contacts, or by an advertiser uploading it to create a custom audience, is used for PII-based advertising. Such use would be particularly pernicious because it involves PII that a user is not even aware Facebook has, and which additionally could be inaccurate (as it is not verified by the user).

Phone contacts One way that Facebook could learn users' PII from their friends would be by scanning friends' contact databases, linking contacts to existing Facebook accounts, and then augmenting the Facebook accounts with any additional PII found in the contacts database. For example, if a Facebook user has a phone contact containing an email address corresponding to some Facebook user, and some phone number that does not correspond to any Facebook user, Facebook might link the new phone number with the account corresponding to the email address.

We used a factory-reset Android phone, and created a contact containing the full name and the email address of one of the authors (both of which Facebook already had), as well as a new phone number that we controlled and had verified was non-targetable. We then installed the Facebook Messenger App, giving it permissions to sync the list of phone contacts. We found that the previously-unused phone number became targetable in 36 days, [13] showing that it had indeed been linked to the corresponding author's account without their knowledge. Making this situation worse, the matched phone number was not listed on the account's profile, nor in the "Download Your Information" archive obtained from Facebook [5]; thus the target user in this scenario was provided no information about or control over how this phone number was used to target them with ads.

Information provided by advertisers As described in Section 2, in order to use PII-based advertising, advertisers first upload lists of PII belonging to customers, and then target the resulting set of Facebook users that match with advertisements. This information is "encrypted" [12] (in reality, hashed) prior to upload. However, because Facebook uses SHA-256 with no salt added, they could potentially determine what PII was uploaded via techniques like rainbow tables. Even without reverse-engineering the uploaded PII, Facebook could potentially use this data to enrich the PII information it uses to match users for targeted advertising as illustrated in the following example. Assume Facebook knows the hashed value h_a for a PII attribute a for a particular user u. Assume an advertiser uploads a record (h_a, h_b) consisting of hashed values for attributes a and b. Using the value of h_a, Facebook can determine that the corresponding user is u, and learn that the hashed value of attribute b for u is h_b; in the future, Facebook can then match h_b to user u without ever knowing the actual value of b.

To study whether Facebook uses the above source, we first upload a list consisting of just a single record containing an email address (which one of the authors uses to log in to Facebook), and another email address that we verified was not targetable. We then check whether the email address becomes targetable; we found that it did *not* become targetable even after a month, suggesting that this source is not used by Facebook to infer PII for advertising.

4.4 Verification by running ads

Out of seven different potential sources of PII studied above, we found that five were indeed used to enable PII-based advertising. Most worrisome, we found that phone numbers uploaded as part of syncing contacts—that were never owned by a user and never listed on their account—were in fact used to enable PII-based adver-

13 This is an upper bound, owing to a short gap in our testing

tising. In all cases we found either no disclosure about the potential uses of the data, or insufficient disclosure in the form of generic statements.

In order to confirm that a given piece of PII has become targetable as a result of us providing it to Facebook through a given source, we run PII-based targeted advertisements targeting each PII found to be targetable (the final step of methodology described in Section 3). As described in that section, this may not be guaranteed to succeed, even if the PII does indeed correspond to the user being targeted (owing to the complexity of the ad delivery process [23]). To increase the chances of our ad winning in the auction process (which is used to decide which ad among a set of competing ads is shown to a user), we use a bid amount four times higher than the default bid shown by Facebook. We then search for our ads by scrolling through the Facebook pages of the corresponding accounts, and identify them by looking for the custom text that we put in each ad.

We were able to successfully target and receive ads targeting the phone numbers added for two-factor authentication, for security notifications, and provided to Facebook as part of uploading the phone contacts; we were not able to successfully target and receive ads for phone numbers added directly to a profile and via Facebook Messenger. Moreover, in the ad campaigns where we were successful, we were only able to receive less than half of the distinct ads targeted towards each. While this result confirms some of our most surprising findings, it also underscores why our methodology for inferring whether a user is targetable is necessary, as relying on placing ads alone is potentially an expensive signal prone to false-negative errors.

5 Discussion

We briefly discuss a few issues that our methodology and findings bring up.

Why use potential reach estimates? At first glance, it would seem that a better and simpler methodology to check whether a source of PII is used for PII-based advertising is: (i) add the PII to a target user's account via the given source, and then (ii) target an ad to an audience with the given PII and check that the target user receives the ad. However, this method has a number of drawbacks. *First*, such an experiment could easily become expensive if Facebook imposed a large minimum size on an audience for an ad to run. *Second*, as previously mentioned, other confounding variables (such as competing ads and Facebook's estimates of ad quality and relevance [23]) might interfere with the results of the experiments; for example, false negatives may arise if the ad launched after adding the PII fails to reach the target users due to competing ads (with better bids, or from more reputed advertisers). *Third*, for sources of PII that the user does not verify (such as phone contacts synced by another user), PII already associated with some other user may not be associated to the target user instead. Therefore, it is essential to be able to check that *no other* user can be targeted with a given PII, which we are able to do by exploiting potential reach estimates using the proposed methodology.

Limitations and challenges Our methodology can only check if data is used for PII-based advertising, and not for advertising in general. Another challenge, general to studying whether a given source of PII is used, is that the service might be running sophisticated algorithms to determine whether or not to use a particular PII for advertising, especially in cases where a user does not directly provide and verify their PII. For example, when using hashed PII records provided by advertisers, Facebook might require a new hashed PII value to occur multiple times in records that match a given user before associating the value with the user. It can thus be challenging to provide PII via a given source in such a way that it passes the checks imposed by any Facebook algorithms. On the flip side, this means that a positive result indicating that Facebook *does* use a given source of PII for PII-based advertising does not mean that Facebook will *always* use any PII provided via that source; further controlled experiments might be necessary to reveal the exact conditions under which Facebook uses PII provided via that source.

Changes to the advertising interface The experiments described in this paper that use the potential reach estimates were all conducted before March 2018. Subsequently, we found that potential reach estimates could be used to leak users' attributes; in response to our disclosure, Facebook removed potential reach estimates for custom audiences [31]. However, when uploading records to create a custom audience, Facebook still provides an estimate of the number of users matched (the *audience size* estimate mentioned in Section 2). Since these estimates also capture the notion of whether a given PII is used by the advertising platform, and were found by prior work [40] to be obfuscated in similar ways (by simple rounding), our methodology could be modified to potentially use these rather than the potential reach. We leave a full exploration to future work.

Generalizability Size estimates are a fundamental feature of any advertising platform as they help advertisers tailor their ad campaigns and plan their ad budgets. Thus, similar methodology to that proposed in this paper can potentially be used across different advertising platforms to study what sources of PII are used for their PII-based advertising features. Besides, since our procedure for reverse-engineering the potential reach estimates in this paper dealt with multiple common ways of obfuscation (noise, rounding etc.), it could potentially illuminate the process of reverse-engineering other size estimates that are obfuscated differently.

6 Related work

We now overview prior work related to this study.

Transparency of targeted advertising Much attention has been dedicated to shedding light on what factors influence the targeting of a particular ad on the web [29, 30, 32, 44] and on specific services [8]. Ad transparency has also been studied in the context of Facebook [2] (with a focus on the explanations that Facebook provides to users as to why they were shown an ad) and Google [41] (showing that Google does not reveal all the categories inferred about a user).

Privacy Because of the lack of information provided by the interfaces where users' data is input, it is often unclear how the PII will be used. For example, Facebook's two-factor authentication interface does not specify the privacy terms that apply to the inserted numbers, nor does it provide the ability to opt out of certain kinds of use. Nonetheless, the company has been using numbers obtained through this interface to send notifications as text messages and to make Friends suggestions as part of its People You May Know feature [19]. Likewise, Facebook has been suspected of using phone numbers collected by users' contact lists to populate the People You May Know feature [24]. Counterintuitively, Tucker [36] shows that giving users control over their privacy can be beneficial for advertising. According to Tucker's study on Facebook, users are more likely to click on an ad if their perception of control over their personal information is higher.

Malicious and discriminatory advertising In 2011, Korolova [27] found that malicious advertisers could infer personally identifying information about users who click on an ad. The attack was based on Facebook's attribute-based 'microtargeting', which has since been disallowed by imposing a minimum audience size of 20. However, subsequent studies showed that targeted advertising platforms are still subject to leaks of personal information [15, 17, 28, 40] and potential discrimination [14, 16]. While Facebook took action to fix these issues [20, 26, 33], the deployment of discriminatory ads was still possible in November 2017 [10]. Speicher et al. [35] demonstrated that the measures taken by Facebook against discrimination (i.e., banning the use of certain attributes, such as 'ethnic affinity') are insufficient. Their work proposes an alternative solution based on a measure for discriminatory targeting that is independent of the attributes used in the ad.

7 Conclusion

Given the incentive advertising platforms have to obtain high-fidelity PII (phone numbers and email addresses) to enhance their services, there is a strong reason to expect the re-purposing of collected PII for targeted advertising. This incentive is exacerbated with the recent introduction of PII-based targeting, which allows advertisers to specify *exactly which users* to target by specifying a list of their PII.

This paper was the first to propose a methodology that uses size estimates to study what sources of PII are used for PII-based targeted advertising. We applied the proposed methodology to investigate which of a potential range of sources of PII were actually used by Facebook for its PII-based targeted advertising platform, confirming that Facebook uses at least five different sources of PII to enable PII-based advertising.

We also examined what is disclosed to users and what controls users have over the PII that is used to target them. We showed that there is often very little disclosure to users, often in the form of generic statements that do not refer to the uses of the particular PII being collected or that it may be used to allow advertisers to target users. Our paper highlights the need to further study the sources of PII used for advertising, and shows that more disclosure and transparency needs to be provided to users.

8 Acknowledgements

We thank the anonymous reviewers for their helpful comments. This research was supported in part by the Data Transparency Lab and NSF grant CNS-1616234.

References

[1] Graph API: Changelog Version 3.0. https://developers.facebook.com/docs/graph-api/changelog/version3.0.

[2] A. Andreou, G. Venkatadri, O. Goga, K. P. Gummadi, P. Loiseau, and A. Mislove. Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations. *NDSS*, 2018.

[3] About Customer Match. https://support.google.com/adwords/answer/6379332?hl=en.

[4] About Potential Reach. https://www.facebook.com/business/help/1665333080167380?helpref=faq_content.

[5] Accessing Your Facebook Data. https://www.facebook.com/help/405183566203254.

[6] Best practices for unique identifiers. https://developer.android.com/training/articles/user-data-ids.

[7] C. Cadwalladr and E. Graham-Harrison. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

[8] A. Datta, M. C. Tschantz, and A. Datta. Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination. *PETS*, 2015.

[9] Facebook. Personal Communication.

[10] Facebook (Still) Letting Housing Advertisers Exclude Users by Race. https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

[11] Facebook Ads Preferences. https://www.facebook.com/ads/preferences.

[12] Facebook Custom Audiences. https://developers.facebook.com/docs/marketing-api/custom-audiences-targeting/v3.1.

[13] Facebook Data Policy. https://www.facebook.com/about/privacy/update.

[14] Facebook Enabled Advertisers to Reach 'Jew Haters'. https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters.

[15] Facebook Leaks Usernames, User IDs, and Personal Details to Advertisers. http://www.benedelman.org/news/052010-1.html.

[16] Facebook Lets Advertisers Exclude Users by Race. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race/.

[17] Facebook Messenger Chatbots Can Leak Your Private Information. https://www.techworm.net/2016/09/facebook-messenger-chatbots-can-leak-private-information.html.

[18] Facebook Response to Questions from Committee on Commerce, Science, and Transportation. https://www.commerce.senate.gov/public/_cache/files/9d8e069d-2670-4530-bcdc-d3a63a8831c4/7C8DE61421D13E86FC6855CC2EA7AEA7.senate-commerce-committee-combined-qfrs-06.11.2018.pdf.

[19] Facebook Turned Its Two-Factor Security 'Feature' Into the Worst Kind of Spam. https://gizmodo.com/facebook-turned-its-two-factor-security-feature-into-th-1823006334.

[20] Facebook adds human reviewers after 'Jew haters' ad scandal. http://www.bbc.com/news/technology-41342642.

[21] Facebook marketing API. https://developers.facebook.com/docs/marketing-apis.

[22] Facebook plans crackdown on ad targeting by email without consent. https://techcrunch.com/2018/03/31/custom-audiences-certification/.

[23] Facebook: About the delivery system: Ad auctions. https://www.facebook.com/business/help/430291176997542.

[24] How Facebook Figures Out Everyone You've Ever Met. https://gizmodo.com/how-facebook-figures-out-everyone-youve-ever-met-1819822691.

[25] How Trump Conquered Facebook—Without Russian Ads. https://www.wired.com/story/how-trump-conquered-facebookwithout-russian-ads/.

[26] Improving Enforcement and Promoting Diversity: Updates to Ads Policies and Tools. http://newsroom.fb.com/news/2017/02/improving-enforcement-and-promoting-diversity-updates-to-ads-policies-and-tools/.

[27] A. Korolova. Privacy Violations Using Microtargeted Ads: A Case Study. *Journal of Privacy and Confidentiality*, 3(1), 2011.

[28] B. Krishnamurthy, K. Naryshkin, and C. E. Wills. Privacy leakage vs. Protection measures: the growing disconnect. *IEEE W2SP*, 2011.

[29] M. Lecuyer, G. Ducoffe, F. Lan, A. Papancea, T. Petsios, R. Spahn, A. Chaintreau, and R. Geambasu. XRay: Enhancing the Web's Transparency with Differential Correlation. *USENIX Security*, 2014.

[30] M. Lecuyer, R. Spahn, Y. Spiliopolous, A. Chaintreau, R. Geambasu, and D. Hsu. Sunlight: Fine-grained Targeting Detection at Scale with Statistical Confidence. *CCS*, 2015.

[31] G. Marvin. Exclusive: Facebook will no longer show audience reach estimates for Custom Audiences after vulnerability detected. 2018. https://marketingland.com/exclusive-facebook-will-no-longer-show-audience-reach-estimates-for-custom-audiences-after-vulnerability-detected-236923/.

[32] J. Parra-Arnau, J. P. Achara, and C. Castelluccia. MyAdChoices: Bringing Transparency and Control to Online Advertising. *ACM TWEB*, 11, 2017.

[33] Protecting Privacy with Referrers. Facebook Engineering's Notes. http://www.facebook.com/notes/facebook-engineering/protecting-privacy-with-referrers/392382738919.

[34] M. Schroepfer. An Update on Our Plans to Restrict Data Access on Facebook. 2018. https://newsroom.fb.com/news/2018/04/restricting-data-access/.

[35] T. Speicher, M. Ali, G. Venkatadri, F. N. Ribeiro, G. Arvanitakis, F. Benevenuto, K. P. Gummadi, P. Loiseau, and A. Mislove. On the Potential for Discrimination in Online Targeted Advertising. *FAT**, 2018.

[36] C. E. Tucker. Social Networks, Personalized Advertising, and Privacy Controls. *Journal of Marketing Research*, 2014.

[37] Target Custom Groups of Twitter Users. https://business.twitter.com/en/targeting/tailored-audiences.html.

[38] The Facebook Business Tools. https://www.facebook.com/help/331509497253087.

[39] The Facebook Companies. https://www.facebook.com/help/111814505650678.

[40] G. Venkatadri, Y. Liu, A. Andreou, O. Goga, P. Loiseau, A. Mislove, and K. P. Gummadi. Privacy Risks with Facebook's PII-based Targeting: Auditing a Data Broker's Advertising Interface. *IEEE S&P*, 2018.

[41] C. E. Wills and C. Tatar. Understanding What They Do with What They Know. *WPES*, 2012.

[42] What are the Facebook Products? https://www.facebook.com/help/1561485474074139.
[43] What's a Custom Audience from a Customer List? https://www.facebook.com/business/help/341425252616329/.
[44] eyeWnder_Experiment. http://www.eyewnder.com/.