On Profile Linkability despite Anonymity in Social Media Systems

Michael Backes
CISPA, Saarland University & MPI-SWS
Saarland Informatics Campus

Pascal Berrang
CISPA, Saarland University
Saarland Informatics Campus

Oana Goga
MPI-SWS
Saarland Informatics Campus

Krishna P. Gummadi
MPI-SWS
Saarland Informatics Campus

Praveen Manoharan
CISPA, Saarland University
Saarland Informatics Campus

ABSTRACT

A number of works have recently shown that the privacy offered by pseudonymous identities on social media systems like Twitter or Reddit is threatened by cross-site identity linking attacks. Such attacks link the identities of the same user across websites. Therefore, assessing linkability, i.e., the risk that identities are linked across different websites, remains an important open problem.

In this work, we analyze whether anonymity within a single social media site can protect a user from being linked across sites. To this end, we first introduce a relative linkability measure ranking identities within a social media site by their anonymity. We show that anonymity alone is not sufficient to assess linkability risks by evaluating this measure on a data set comprising 15 million comments gathered from the Reddit social media system.

Second, we mitigate this insufficiency and present our absolute linkability measure, which, in addition, utilizes information about matching identities. Then, we confirm the validity of this measure on our data set. The measure is able to accurately assess the linkability risk in almost 75% of the cases and, more importantly, is shown to never underestimate the linkability risk.

1. INTRODUCTION

Social media systems, where any user can join the system and contribute content, are becoming widely popular. Examples of social media systems include blogging sites like Twitter and LiveJournal, social bookmarking sites like Delicious and Reddit, and peer-opinion sites like Yelp, Amazon, and eBay reviews. To enable users to contribute freely and without fear, these sites need to offer their users *anonymity*. Today, many systems allow users to operate using pseudonymous identities that can be created without providing any certification by trusted authorities and where users determine what information they choose to reveal about themselves. For instance, many Twitter users do not provide (or deliberately provide fake) information about their real names, bios, or profile photos when creating identities.

Many users participate in different social media sites assuming different pseudonymous identities under the belief that their identities across different sites cannot be linked. However, recently researchers have shown that adversaries can exploit seemingly innocuous and latent information such as location patterns [11] and linguistic patterns in *public posts* [18] to link even pseudonymous identities that a user has created across different sites. Such attempts to aggregate and link user data across multiple social media sites in order to reveal a more comprehensive profile of the information sources have many commercial applications [2], but they also raise serious privacy concerns for the users of these sites.

Contributions In this paper, we examine the degree to which the anonymity of a user's identity can be used to estimate the linkability threats that are *inherent* to the *publicly visible content* contributed by a user to social media sites. That is, we evaluate linkability threats assuming that the only data that is available for linking a user's identities are the contents of the public posts written using the identities. In practice, an adversary might have additional data about a user (e.g., non-public data such as a user's IP address or a user's real name) that might help them link the user's identities. However, we consider only public posts of the user as (i) they are available to all adversaries and (ii) they represent the minimum amount of information a user reveals by participating in the social media site. Consequently, we consider the unavoidable linkability threat that arises from a user's content contributions to different social media sites.

Our work is motivated by the relation of linkability and anonymity of a user's identities in a traditional database setting. In such a setting, anonymity usually requires equality within an anonymity set, which naturally implies unlinkability of the user's identities. The same, however, cannot directly be applied to the linkability of user posts in social media systems like Facebook, Twitter or Reddit since, on such platforms, information is presented in a highly unstructured manner: traditional privacy models, such as k-anonymity [24], l-diversity [17], t-closeness [16], or differential privacy [9], have been defined over well-structured databases and cannot be applied to user posts (e.g., it is not clear what the quasi-identifiers and sensitive attributes

WPES'16, October 24 2016, Vienna, Austria

ISBN 978-1-4503-4569-9/16/10...$15.00

DOI: http://dx.doi.org/10.1145/2994620.2994629

in this context are). Moreover, it is unclear how differentially private noising would work on natural language posts.

We leverage recent work that extends the notion of k-anonymity over structured data sets to unstructured data sets [6]: For a user identity u in a social media system, (k, d)-anonymity captures the largest k subset of identities containing u such that every identity within the subset is within a divergence (or dissimilarity) threshold of d from u.

Using (k, d)-anonymity, we evaluate whether anonymity in one social media system allows us to estimate the risk of linkability threats across social media systems. Specifically, we address the following two questions: (i) Can the knowledge of (k, d)-anonymity of users in an online social media system be used to estimate their *relative linkability risks*, i.e., estimate whether one user is more at risk of her identities being linked than other users? (ii) To what extent does combining knowledge of (k, d)-anonymity of a user in a social media system with information about their matching identity in a different social media system improve the linkability assessment?

We use an extensive data set of over 15 million comments posted by users across 1,930 topical communities in the Reddit social media system. Using potential strategies of a rational adversary, we analyze the correlations between the (k, d)-anonymity of a user's identity and the estimated risk of the identity being matched to determine the utility of the (k, d)-anonymity measure.

Our findings yield several valuable insights about the relation between anonymity and linkability. First, the ranking of identities by the size of their (k, d)-anonymity set positively correlates with the matching set size (i.e., the number of identities the adversary considers as potentially matching). However, this correlation is fairly weak and we thus conclude that *anonymity alone is not sufficient to assess linkability risks on social media systems.* Second, extending (k, d)-anonymity with information about the matching identities yields a more useful linkability risk assessment. Using the local matching set μ that we derive by combining anonymity sets and information about matching identities we can successfully estimate the size of the matching set: in over 74% of the cases, the size of the local matching set μ is at least 0.8 times the actual matching set size of the adversary.

Outline We begin by introducing required background knowledge and motivating our work in Section 2. We then develop the relative and absolute linkability measures in Section 3. In Section 4, we introduce the Reddit data set we use for our evaluations. Using this data set, we then evaluate, in Section 5, both linkability measures and show that anonymity alone is not a good measure of linkability, but extending anonymity with information about matching identities can provide a good measure of linkability. We elaborate on related work in Section 6 before concluding in Section 7.

2. BACKGROUND AND MOTIVATION

Before examining how well anonymity can be used to assess linkability threats that allow an adversary to link identities across sites, we first have to discuss the terminology we use in the remainder of the paper and provide the background on key concepts that underlie our work.

2.1 Domains and Identities

The term *identity* denotes the profile created by a user in a social media system. A *domain* is the collection of identities within a social media system. A pair of identities within different domains is called *matching* if they belong to the same user.

2.2 Identity Representation and Similarity

The first challenge in addressing the anonymity and linkability threats in social media systems is to find a suitable representation of identities. Given that the *only* information that we presume to know about an identity are its public posts, we represent each identity by fitting a statistical model to the identity's textual posts.

The simplest way to construct such a statistical model would be to determine the relative frequency of each word unigram used by an identity. Specifically, we represent identities through a unigram-statistical language model that captures the relative frequency with which the identity uses a specific unigram w: i.e., given a vocabulary $\mathcal{V}$ of word unigrams and the collection of comments $C_{\mathcal{I}}$ by $\mathcal{I}$, the identity model $\theta_{\mathcal{I}}$ is defined by

$$Pr[w \mid \theta_{\mathcal{I}}] = \frac{\mathsf{count}(w, C_{\mathcal{I}})}{\sum_{w' \in \mathcal{V}} \mathsf{count}(w', C_{\mathcal{I}})}.$$

While this identity model is fairly simple, it is sufficient to assess the relation between anonymity and linkability in social media systems that allow the sharing of user-generated text content. In Section 5, we investigate various more complex models. The general observations, however, stayed the same. It would also be possible to incorporate other sources of information – as for example pictures, videos or location – into the identity model. Naturally, the precise anonymity and linkability risk of an identity will then change with such an extended model that includes a wider variety of features, however, in this paper we are rather interested in gaining conceptual clarity into the ways anonymity and linkability relate to each other, rather than estimating the precise linkability risks of an identity in a specific system and under specific scenarios.

To measure how *similar* two identities are we use the Jensen-Shannon divergence [10] D_{JS}.[1] The Jensen-Shannon divergence is a symmetric variant of the popular Kullback-Leibler divergence, which has been used with large success to determine the similarity of probability distributions (and therefore statistical models), and the square root of D_{JS} provides a full-fledged metric. In the remainder of the paper we will talk about the *distance* $\mathsf{dist}(\mathcal{I}, \mathcal{I}') = \sqrt{D_{\mathsf{JS}}(\theta_{\mathcal{I}}, \theta_{\mathcal{I}'})}$ of identities, induced by this divergence measure, instead of their similarity, to provide a better, intuitive understanding.

2.3 Adversarial Matching Strategy

In this paper, we consider an adversary that tries to link a source identity $\mathcal{I}_{\mathcal{S}}$ within a source domain $\mathcal{S}$ to the matching target identity $\mathcal{I}_{\mathcal{T}}$ within a target domain $\mathcal{T}$. We assume that the adversary has at her disposal (i) the posts of all identities in both domains and (ii) a small ground-truth set of matching identities across both domains. These are standard assumptions made by the majority of previous work in this area [12].

[1]We also tested other metrics such as Cosine similarity in Section 5.1, but the results were not affected significantly.

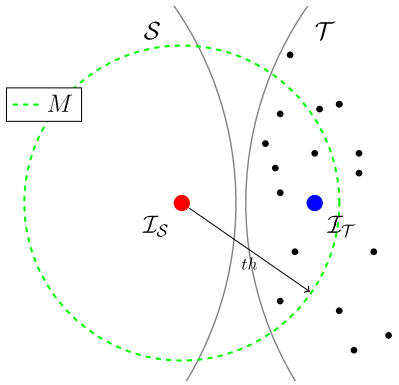


Figure 1: Illustration of two domains and the matching set $\mathcal{M}$ of $\mathcal{I}_\mathcal{S}$ in $\mathcal{T}$. The size of the matching set is 8.

The matching process of the adversary Adv consists of four steps: (*i*) Adv computes the pairwise similarity between all identities in $\mathcal{S}$ and all identities in $\mathcal{T}$; (*ii*) he computes the likelihood of any two identities to belong to the same user based on their similarity[2]; (*iii*) he then ranks all pairs of identities according to their likelihood of belonging to the same user; and (*iv*) the adversary chooses a threshold th on the likelihood measure (according to how accurately he wants to link identities) and links all the identities that are above the threshold. The threshold choice is the standard trade-off between recall (i.e., the fraction of identities linked out of all matching identities) and precision (i.e., the probability that the identities linked are actually matching identities) calculated over the ground-truth set of matching identities. This strategy is consistent with the strategy employed by the majority of previous works on matching identities. We discuss several such works in Section 6.

While the matching strategy we consider in this paper corresponds to a rational adversary, who wants to increase the number of identities he can link correctly, this adversary model does not necessarily represent the worst case adversary; and an adversary could simply choose to not be rational. As pointed out by Backes et al. [6] it is, in general, impossible to provide unlinkability guarantees against arbitrary adversaries in open and unstructured settings that we consider in this work.

2.4 Linkability of Identities

Through his choice of the threshold value th (see Section 2.3) the adversary defines the set of identities within the target domain $\mathcal{T}$ that he considers *potentially matching* the source identity $\mathcal{I}_\mathcal{S}$: we call this set the *matching set* $\mathcal{M}(th)$ of the adversary for identity $\mathcal{I}_\mathcal{S}$. We illustrate such a matching set in Figure 1. The matching set is the set of identities from which the adversary cannot sufficiently distinguish which target identity $\mathcal{I}_\mathcal{T}$ (cf. Figure 1) is related to $\mathcal{I}_\mathcal{S}$.

We can therefore quantify the linkability of a user's identities using this matching set: the bigger $\mathcal{M}(th)$ is, the less likely it is that the adversary will link $\mathcal{I}_\mathcal{S}$ to $\mathcal{I}_\mathcal{T}$. Note that the size of the matching set of an identity depends on the threshold th chosen by the adversary. In this paper, we will consider both scenarios where we know and where we do not know the adversary's threshold choice when estimating the linkability risks of identities.

[2]An adversary can consider the similarity between two identities ($\mathcal{I}_\mathcal{S}$ and $\mathcal{I}_\mathcal{T}$) as the likelihood of them to belong to the same user, or he can compute more complex functions that, in addition to the similarity between $\mathcal{I}_\mathcal{S}$ and $\mathcal{I}_\mathcal{T}$, take into account the similarity between $\mathcal{I}_\mathcal{S}$ and other identities in $\mathcal{T}$.

2.5 Anonymity of an Identity

We formalize anonymity in a social media system using the notion of (k, d)-anonymity, recently put forward by Backes et al. [6]. At a high level, the notion of (k, d)-anonymity provides a generalization of the classic notion of k-anonymity [24]: (k, d)-anonymity defines the *anonymity set* $\mathcal{A}(d)$ of the target identity $\mathcal{I}_\mathcal{T}$ that contains at least k identities within the target domain $\mathcal{T}$ that have a distance of at most d to $\mathcal{I}_\mathcal{T}$.

Definition 1 ((k, d)-Anonymity). *An identity $\mathcal{I}$ is* (k, d)-anonymous *in a domain $\mathcal{D}$ if there exists an* anonymity set $\mathcal{D}' \subseteq \mathcal{D}$ *with the properties that $\mathcal{I} \in \mathcal{D}$, that $|\mathcal{D}'| \geq k$ and that all $\mathcal{I}' \in \mathcal{D}$ have $\mathsf{dist}(\theta_\mathcal{I}, \theta_{\mathcal{I}'}) \leq d$.*

We denote with $\mathcal{A}_\mathcal{I}(d)$ the largest anonymity set of $\mathcal{I}$ for a distance of d, and call d its convergence.

Throughout the remainder of the paper we forgo the subscript of the anonymity set $\mathcal{A}(d)$ when we talk about the anonymity set of the target identity $\mathcal{I}_\mathcal{T}$ to keep the notation simple.

2.6 Relation of Anonymity and Linkability

In the traditional database setting, anonymity naturally implies unlinkability: notions such as k-anonymity and l-diversity require all identities within an anonymity set to be equivalent. Thus, any source identity cannot be uniquely linked to any target identity in a sufficiently large anonymity set. Ideally, we would want the same to hold in open settings such as social media systems as well: if an identity $\mathcal{I}_\mathcal{T}$ is anonymous in its domain $\mathcal{T}$, it should also be difficult to link it to its matching identity $\mathcal{I}_\mathcal{S}$ since the adversary cannot sufficiently distinguish $\mathcal{I}_\mathcal{T}$ from the other identities in $\mathcal{T}$.

The main question we pose in this paper is whether *the anonymity set size of the target identity $\mathcal{I}_\mathcal{T}$ provides a good assessment of the difficulty of successfully linking the source identity $\mathcal{I}_\mathcal{S}$ to $\mathcal{I}_\mathcal{T}$, i.e., does a large anonymity set imply a large matching set?* Using the notions we introduced in the previous section, our goal is therefore to investigate whether the $\mathcal{I}_\mathcal{T}$'s anonymity, as estimated by its (k, d)-anonymity, can be used to estimate the size of the $\mathcal{I}_\mathcal{S}$'s matching set $\mathcal{M}$.

3. ASSESSING LINKABILITY RISKS USING ANONYMITY

We investigate two different scenarios in which we use an identity's anonymity to assess its linkability across social media systems. In the first scenario, we assume that we do not know the adversary's matching strategy (i.e., we do not know the threshold he chooses to link identities – see Section 2.3) and we do not know the matching identities of users in other social media systems. Our goal is to see whether the relative anonymity of identities in the social media system can be used to derive a relative linkability measure that informs the users about their linkability risks. In the second scenario we assume the attacker is targeting a particular user and hence, we can combine the (k, d)-anonymity of the $\mathcal{I}_\mathcal{T}$ as well as knowledge about its matching identity $\mathcal{I}_\mathcal{S}$ to develop an *absolute linkability measure*.

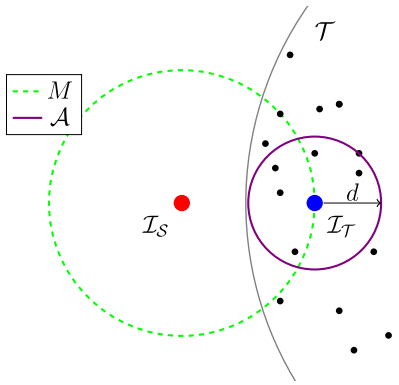


Figure 2: Illustration of the anonymity set $\mathcal{A}$ and the matching set $\mathcal{M}$.

Consistent_Ranking($\mathcal{T}$, $\mathbb{D}$)

for $d \in \mathbb{D}$ **do**
  for $\mathcal{I} \in \mathcal{T}$ **do**
    compute $\mathcal{A}_\mathcal{I}(d)$
  sort all $\mathcal{A}_\mathcal{I}(d)$ into list $\mathcal{L}$
  for $\mathcal{I} \in \mathcal{D}$ **do**
    $\mathsf{rank}_d(\mathcal{I}) = \mathsf{fillingCompRank}(\mathcal{A}_\mathcal{I}, \mathcal{L})$
$G = (V = \mathcal{T} \cup \{1, .., |\mathcal{T}|\}, E = \mathcal{T} \times \{1, .., |\mathcal{T}|\}, w)$ with $\forall e \in E : w(e) = 0$
for $d \in \mathbb{D}$ **do**
  for $\mathcal{I} \in \mathcal{T}$ **do**
    $w((\mathcal{I}, \mathsf{rank}_d(\mathcal{I})))+ = 1$
compute maximum weight matching M on G
for $(\mathcal{I}, r) \in M$ **do**
  $\mathsf{rank}^*(\mathcal{I}) = r$
return rank^*

Figure 3: Consistent Ranking of Identities.

3.1 Relative Linkability Measure

Context With the relative linkability measure, we want to identify those identities within a domain that are most susceptible to being successfully linked to their matching identities in other domains. Intuitively, and without knowledge about the matching identity, this mostly depends on the uniqueness of an identity within a domain: observe within the same domain that an identity either (a) is very unique and therefore easily identifiable, or (b) blends well into the crowd and therefore has good anonymity.

The notion of (k, d)-anonymity we introduced in Section 2.5 essentially captures the uniqueness of the target identity $\mathcal{I}_\mathcal{T}$ in the target domain $\mathcal{T}$. Our hope is that by ranking identities by their anonymity sets, we get a relative assessment of an identity's linkability compared to other identities within the same domain. Against a rational adversary that tries to maximize the number of correct matchings he achieves between two domains, such a relative ranking provides insight into which identity is more likely to be matched first by the adversary.

Since (k, d)-anonymity has two parameters, we have two options to generate a suitable ranking to predict the relative linkability of user within a domain. The first is to rank identities by their anonymity set size: for a given convergence value d, we compute, for all identities $\mathcal{I} \in \mathcal{T}$, the anonymity set $\mathcal{A}(d)$ and rank the identities by its size. The second option is to rank identities by the convergence of their anonymity sets: here, we fix the anonymity set size k and determine the required convergence value d to achieve k. The identities are then ranked by Independent of how we approach this ranking, the linkability assessment of a specific identity is then derived from its rank: the relative linkability measure thus tells each identity how linkable it is compared to other identities in the same domain.

However, at this point, we do not have any additional information that would support the choice of any specific value for d or k. Instead, we propose a ranking scheme that combines the rankings computed for multiple values of d or k to generate an overall consistent ranking. In the following, we describe this consistent ranking scheme for ranking by anonymity set size. The algorithms can be easily adopted similarly for ranking by convergence.

Consistent Ranking of Identities Given a set of convergence values $\mathbb{D}$ (in our evaluation, we choose all convergence values between 0 and 1, in $\frac{1}{1000}$ steps, since the Jensen-Shannon divergence is bounded by these values), we compute for all identities $\mathcal{I} \in \mathcal{T}$ and for all convergences $d \in \mathbb{D}$ the maximum anonymity set $\mathcal{A}_\mathcal{I}(d)$ and rank each identity by the size of these anonymity sets in rank_d. During this ranking, we resolve ties by assigning all identities that have equal set sizes the set of ranks they could occupy. For example, if rank 3 and 4 are not uniquely defined because of a tie between two identities, both will be assigned the set of ranks $\{3, 4\}$. This procedure that we call $\mathsf{fillingCompRank}$ corresponds to a standard competition ranking with filling up the gaps afterwards.

Next, we construct a bipartite graph $G = (V = \mathcal{T} \cup \{1, .., |\mathcal{T}|\}, E = \mathcal{T} \times \{1, .., |\mathcal{T}|\}, w)$ between all identities and their rankings. The weight of an edge $(\mathcal{I}, r)$ in the bipartite graph corresponds to the number of times $\mathcal{I}$ was ranked at rth position in the rank_d rankings.

The final ranking is then determined by the maximum weight matching on the bipartite graph. This ranking scheme takes into account how large the anonymity sets of identities are and also how quickly they grow for varying values of d. A pseudo-code implementation of this algorithm can be found in Figure 3.

In the experimental evaluation in Section 5.4, we evaluate this consistent ranking method in practice.

3.2 Absolute Linkability Measure

Context Contrary to the relative linkability measure, we now make additional assumptions about the adversary: we consider a different scenario in which we know which matching identities $\mathcal{I}_\mathcal{S}$ and $\mathcal{I}_\mathcal{T}$ the adversary wants to link. For the absolute linkability measure, we include additional information about the source identity $\mathcal{I}_\mathcal{S}$ to produce a targeted estimate of linkability. Our goal is to estimate how many identities in the target domain $\mathcal{T}$ match the source identity $\mathcal{I}_\mathcal{T}$ at least as well as the matching target identity $\mathcal{I}_\mathcal{T}$, i.e., we want to estimate the size of the matching set $\mathcal{M}$.

A first, simple approach to include information about the source identity $\mathcal{I}_\mathcal{S}$ in our linkability assessment is to choose the convergence d of the anonymity sets $\mathcal{A}(d)$ as the distance of source and target identity, i.e., $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$.

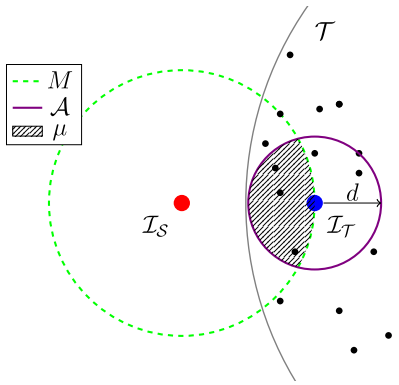


Figure 4: Illustration of the anonymity set $\mathcal{A}$, the matching set $\mathcal{M}$, and the local matching set μ.

Through this, we capture all identities in the neighborhood of $\mathcal{I}_\mathcal{T}$ that can potentially appear in the matching set. While other identities, which are not in $\mathcal{A}(d)$, will still appear in the matching set, considering $\mathcal{A}(d)$ might potentially allow us to provide a lower bound estimate on the size of the matching set. However, in some cases, the anonymity set $\mathcal{A}(d)$ will not approximate the size of the matching set $\mathcal{M}$ well: $\mathcal{A}(d)$ might be distributed in such a way that all identities within $\mathcal{A}(d)$ have a distance $d' \geq d$ to the source identity $\mathcal{I}_\mathcal{S}$, and thus $\mathcal{M} \cap \mathcal{A}(d) = \emptyset$. In the illustration in Figure 4, this would correspond to the hypothetical case where all identities within $\mathcal{A}$ are outside the matching set $\mathcal{M}$.

Therefore, instead of directly estimating the size of the matching set $\mathcal{M}$ with the anonymity set $\mathcal{A}(d)$, we use the *local matching set* μ, which is the intersection between $\mathcal{M}$ and $\mathcal{A}(d)$ to estimate the size of $\mathcal{M}$.

Definition 2 (Local Matching Set). *Let $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$. Then the* local matching set μ *of the source identity $\mathcal{I}_\mathcal{S}$ matching against a target identity $\mathcal{I}_\mathcal{T}$ is defined by $\mu = \mathcal{M} \cap \mathcal{A}(d)$.*

We illustrate the relation between the matching set $\mathcal{M}$, the anonymity set $\mathcal{A}(d)$ and the local matching set μ in Figure 4. Setting the convergence d of the anonymity set to the distance of the matching identities allows us to capture a large part of the identities from the matching set in our local matching set.

4. REDDIT DATA SET

We use Reddit [1] to study the relationship between anonymity and linkability in social media systems. Reddit was founded in 2005 and is one of the largest discussion and information sharing platforms in use today. On Reddit, users share and discuss topics in a vast array of topical *subreddits* collecting all topics belonging to one general area; e.g. there are subreddits for world news, TV series, sports, food, gaming, and many others. Each subreddit contains so-called *submissions*, i.e., user-generated content that can be commented on by other users.

For our evaluation, we use the data set collected in [6]. The data set was collected during September 2014 through Reddit's own API. The data set contains more than 40 million comments spanning over 44,000 subreddits. The comments were written by about 81,000 different users.

4.1 Data on Matching Identities

Since we aim to asses the risk of linking the identities of the same user across different communities, it is crucial to have ground-truth on matching identities. We opportunistically use Reddit's subreddit structure to obtain such ground-truth: we treat each subreddit as its own (virtual) domain, and assume that each user has a separate identity in each subreddit. This way, we easily obtain the ground-truth on matching identities, because each user has the same pseudonym across all subreddits. Overall, our data set contains about 2.75 million of such identities.

4.2 Ethical Concerns

For our evaluation, we only collected publicly available, user-generated text content from the social media system Reddit and replaced the pseudonyms under which this content was posted with randomized, numerical identifiers. In our evaluation, we did not infer any further information about the users; in particular we did not directly link any profiles, but used the pseudonym information to match the same user's content across different subreddits. We thus do not infer any further sensitive information (through linking) than what is already publicly made available by each user on the Reddit platform.

Since our institutes do not have an IRB, we consulted the opinion of a local privacy lawyer, who confirmed that our research is in accordance with the Max Planck Society's ethics guidelines as well as with the applicable German data protection legislation (§28 BDSG).

4.3 Data Processing

Filtering Identities To avoid noise due to the lack of data, as in [6], we perform our evaluation only on identities that have at least 100 comments and that belong to a subreddit with at least 100 profiles. Through this, we make sure that (a) each identity provides a sufficient amount of comments to model them (a similar approach has been taken in previous work on author identification as well [18]) and (b) there are sufficient identities within a domain to analyze the distribution of anonymity sets. Furthermore, we dropped the three largest subreddits from our data set to speed up the computation.

After filtering, we have a data set that contains 15 million comments contributed by 58,091 different identities that belong to 37,935 different users in 1,930 different subreddits. Details about the distribution of identities over the subreddits can be found in the supplementary material [5].

Normalizing Comments To get a clean representation of the comments to apply the statistical identity models on, we consider the comments after the application of several normalization steps, as described in [6]. These normalization steps include converting comments into lower case, removing Reddit formatting, replacing URLs by their host names, and filtering out stop words.

5. REDDIT EVALUATION

In this section, we evaluate the utility of (k, d)-anonymity to assess the risk of user's identities to be linked across social media systems. We first characterize the size of matching sets and anonymity sets in our Reddit data set. We then investigate whether the relative and absolute linkability mea-

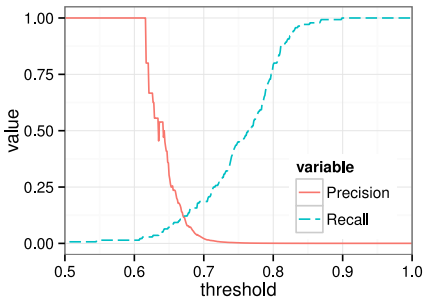


Figure 5: Precision and recall tradeoff for matching identities from subreddit *news* to *worldnews*.

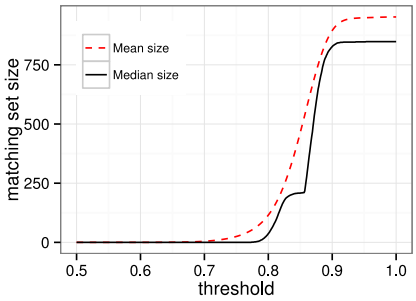


Figure 6: Median and mean matching set sizes of the adversary depending on the chosen threshold (for matching identities from subreddit *news* to *worldnews*).

sures we proposed in Section 3 are a good estimator for the linkability risks, i.e., the matching set size of identities.

Note that, for simplicity, all graphs in this section are based on the source subreddit *news* and the target subreddit *worldnews* if not explicitly mentioned otherwise. During our evaluation process, we also considered other pairs of subreddits for which we provide the same kind of diagrams as supplementary material [5]. For each claim, we also provide general graphs summarizing over the whole data set and showing that the results hold across other subreddits as well.

5.1 Identity Model Instantiations

As mentioned in Section 2.2, we evaluated our measures using various other identity model instantiations. More precisely, we instantiated the identity models not only using (1) unigram frequencies, but also using (2) unigram based indicator vectors, (3) term frequency-inverse document frequencies (TFIDFs), and (4) disjoint author-document topic models [23]. While the first two instantiations do not incorporate the distribution of words within a subreddit, the latter two instantiations were specifically used to separate words belonging to the general topic of a subreddit from author specific language.

For each of these instantiations choices, we evaluated both, the relative and the absolute linkability measures using two different distance/similarity metrics, namely the (a) Jensen-Shannon divergence and (b) Cosine similarity. Our experiments showed that the choice of the distance metric mainly results in a shift of the similarities (and hence a shift of the thresholds) without affecting the precision, the recall and the general take-aways. Moreover, while the conclusions drawn from the experiments remained the same for all instantiations of the identity models, the unigram frequency and the TFIDF approaches provided the best, albeit very similar, results with respect to both our estimations and the adversary's linkage attack. Thus, we will, in the following, focus on the results obtained using unigram frequencies for our identity model and the Jensen-Shannon divergence as our similarity metric.

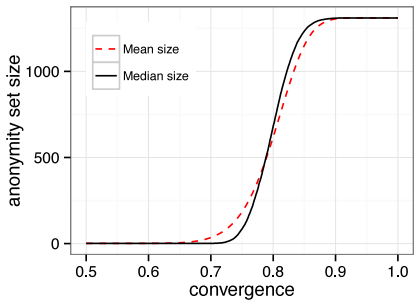


Figure 7: Median and mean anonymity set sizes when varying the convergence d (subreddit *worldnews*).

5.2 Characterization of Matching Sets

In Section 2.3, we explained that a rational adversary tries to correctly match as many identities as possible. To this end, the adversary needs to choose an appropriate threshold on the likelihood that two identities belong to the same user to consider two identities as matching. If the adversary has access to a small ground-truth set (which is the assumption that many previous works in this area make) then he can choose the threshold by analyzing the tradeoff between precision (how many of the identities linked are true matching identities) and recall (how many identities are linked out of all true matching identities). In this paper, we assume that the adversary takes the distance between identities as the likelihood measure. Figure 5 depicts both the precision and recall of an adversary for varying thresholds for matching identities in the *news* subreddit to identities in the *worldnews* subreddit.

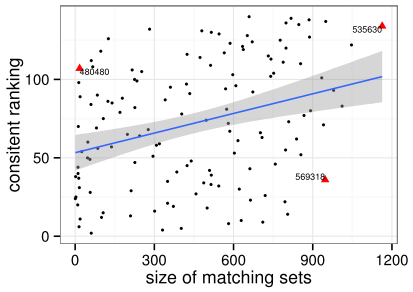
Since the choice of threshold will of course impact the size of the matching sets we plot, in Figure 6, the median and mean size of the matching sets depending on the threshold. For example, the median matching set size for a threshold of 0.8 is 37.

5.3 Characterization of Anonymity Sets

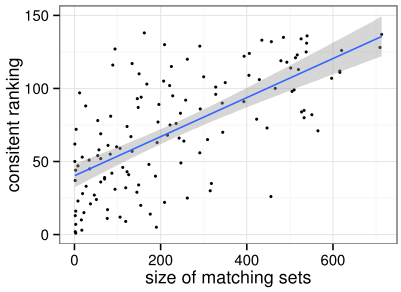
Since the notion of anonymity sets lays the foundation of our two linkability measures, we first have a closer look at its characteristics in our data set. Figure 7, plots the median and mean size of the anonymity set (for the subreddit *worldnews*) depending on the convergence d. For example, the median anonymity set size for a convergence of 0.8 is 37, which is promising as it is similar with the median matching set size for the same threshold.

5.4 Assessing the Relative Linkability Measure

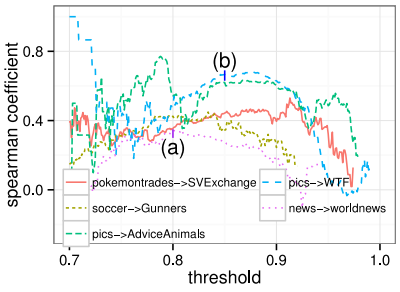
Remember that, in Section 3.1, we introduced the relative linkability measure to identify, within a domain, the identi-



(a) The consistent ranking (over all convergences) compared to the size of the corresponding matching sets for an adversary's threshold of 0.8 (subreddit *news* to *worldnews*).

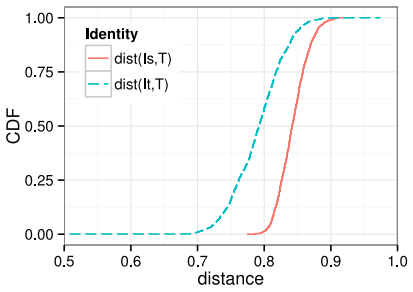


(b) The consistent ranking (over all convergences) compared to the size of the corresponding matching sets for an adversary's threshold of 0.85 (subreddit *pics* to *wtf*).

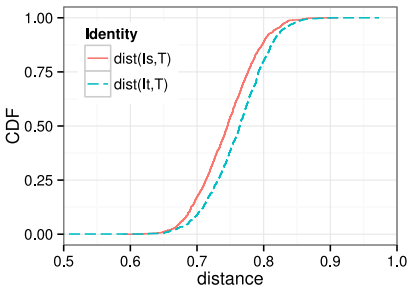


(c) Spearman's correlation coefficient between the consistent ranking and the size of the matching sets for different adversary thresholds and different subreddits.

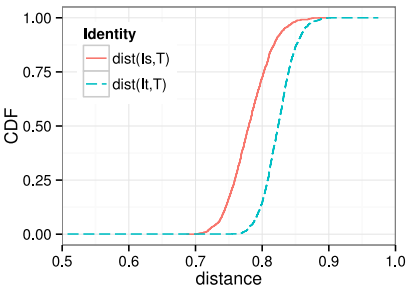
Figure 8: The relative linkability measure.



(a) Case of underestimation of the linkability risk (for user id 480480).



(b) Case of good estimation of the linkability risk (for user id 535630).



(c) Case of overestimation of the linkability risk (for user id 569318).

Figure 9: CDFs of distances from $\mathcal{I}_\mathcal{T}$ and $\mathcal{I}_\mathcal{S}$ to the target subreddit $\mathcal{T}$.

ties that are most at risk of being linked to their matching identities in other domains. In this section, we investigate whether the relative linkability measure is a good estimate of linkability risks. Thus, our goal is to investigate to which degree the consistent ranking provided by the relative linkability measure correlates with the matching set size $\mathcal{M}(th)$.

Note that since this measures relies only on a minimal amount of information, i.e., it only takes into account the similarities between identities in a single domain and does not take into account the matching identities of a user, we do expect the approximation not to hold in all cases.

Figures 8a and 8b depict the correlation between the size of an identity's matching set (for a particular threshold) and the rank of the identity for two different pairs of subreddits. In both figures, we see a positive correlation between the consistent ranking and the size of the matching set, however, Figure 8b presents a better correlation than Figure 8a.

To illustrate how the correlation depends on the threshold chosen by the adversary and the pairs of subreddits considered, Figure 8c depicts the Spearman correlation coefficient between the consistent ranking and the size of the matching set for various thresholds and multiple subreddit pairs. For reference, the thresholds for the previous figures are also annotated. The figure shows that there is a positive correlation between the consistent ranking and the size of the matching set for other pairs of subreddits as well. However, for all the thresholds considered, the correlation is not very strong in general.

Furthermore, in Figure 8a we can see that there are many points that are far from the regression model. There are identities with a high rank that have a small matching set, and there are identities with a low rank that have comparatively large matching sets. While the consistent ranking overestimates the linkability risk of the identity in the bottom right corner which might not be so problematic; it underestimates the linkability risk for the identity in the top left corner which is really problematic because it makes the identity subject to a false sense of anonymity.

To investigate why in some cases the consistent ranking estimates well the size of the matching set while in other cases it overestimate or underestimates it, we investigate further the three highlighted identities. In Figure 9, we analyze the relation between the distances $\mathsf{dist}(\mathcal{I}_T, \mathcal{T})$ from the target identity $\mathcal{I}_T$ to the target subreddit $\mathcal{T}$ and the distances $\mathsf{dist}(\mathcal{I}_S, \mathcal{T})$ from the source identity $\mathcal{I}_S$ to our target subreddit $\mathcal{T}$. To this end, we present the CDFs of these distances for the three identities that have been highlighted in the previous figure (a case of underestimation, a case of good estimation and a case of overestimation). In both Figure 9a (underestimation) and Figure 9c (overestimation), the distributions are rather dissimilar while in Figure 9c (good estimation) the distributions are rather similar. In the first case, the distances to identities in the same subreddit (from

$\mathcal{I}_T$ to $\mathcal{T}$) are smaller than those when matching from the outside (from $\mathcal{I}_T$ to $\mathcal{T}$) which leads to large anonymity sets and small matching sets, which leads in turn to the false sense of anonymity for that particular identity. In the second case, the distances to identities in the same subreddit are larger, which consequently leads to an overestimation of the identity's linkability risk. Thus, the accuracy of the (k, d)-anonymity to estimate the linkability risk depends on how an identity $\mathcal{I}_T$ is placed with respect to other identities in the domain (as measured by the similarity between them) as well as on how far the matching identity $\mathcal{I}_S$ is from identities in the target domain. The absolute linkability measure takes exactly this into account.

5.5 Assessing the Absolute Linkability Measure

The absolute linkability measure, as explained in Section 3.2, aims to assess the linkability risk if an adversary targets a particular user. Thus, the goal of this section is to investigate whether the anonymity set $\mathcal{A}(d)$ where $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$ and the corresponding local matching set μ estimate reliably the size of the matching set $\mathcal{M}(th)$ for a threshold $th = d$.

Anonymity Set Figures 10 depict the correlation between the size of the anonymity set $\mathcal{A}(d)$ and the size of the matching set $\mathcal{M}(d)$ for matching identities between subreddits *news* and *world news*. Note that, compared with the previous section where we had the same th for all pairs of identities, here, for each pair or identities we compute $\mathcal{A}(d)$ and $\mathcal{M}(d)$ where $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$. The Spearman's correlation coefficient for this plot is 0.41, comparable to the values we obtained in the previous section.

To check the general correlation of anonymity sets and matching sets in other subreddits, Figure 11 inspects the ratio of anonymity set sizes and matching set sizes $\frac{|\mathcal{A}(d)|}{|\mathcal{M}(d)|}$ on our whole data set. When over-approximating the risk of an identity the anonymity set size is small compared to the size of the matching set, resulting in a ratio < 1. Conversely, when under-approximating the risk, the ratio is > 1. Note that the x-axis is plotted in log scale to allow us to display the tail ends of the distribution. We can see a clear peak in the area around 1: for at least 71% of all cases, the fraction $\frac{|\mathcal{A}(d)|}{|\mathcal{M}(d)|}$ lies in the interval $[0.8, 1.2)$. Thus, for most pairs of subreddits there is correlation between $\mathcal{A}(d)$ and $\mathcal{M}(d)$. However, the anonymity set size suffers from the same drawback as the relative linkability measures, it underestimates the linkability risk for 41.2% of identities in our data set which makes the anonymity set size not a reliable measure of linkability risks.

Local Matching Set Figure 12a depicts the size of an identity's local matching set compared to the size of the adversary's matching set for our exemplary pair of subreddits. Clearly, except for a few outliers, both sizes positively correlate. The few outliers only provide an over-approximation of the identity's risk: While the local matching set is small and the identity does not seem to blend into the crowd, the matching set is large and thus the identity cannot easily be linked.

The more intriguing question, however, is how accurately the local matching set estimates the matching set. To this end, we analyze the ratio $\frac{|\mu|}{|\mathcal{M}|}$ between both sizes on our whole data set in Figure 12b. If both set sizes coincide, the ratio yields 1, whereas inaccurate estimations of the matching set size will result in a ratio towards 0. We can see that for the vast majority of identities, both sets coincide or at least are very similar: In at least 74% of the cases, the local matching set has at least $0.8 \cdot |\mathcal{M}(d)|$ elements.

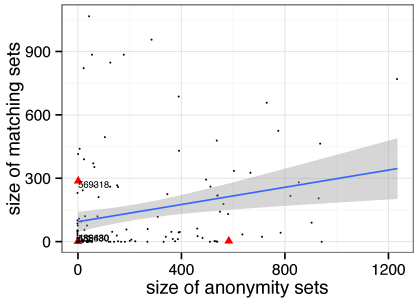


Figure 10: Size of the anonymity set $\mathcal{A}(d)$ compared to the size of the matching set $\mathcal{M}(d)$ where $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$ for each pair of identities (subreddit *news* to *worldnews*).

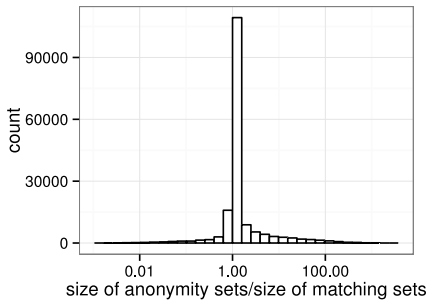
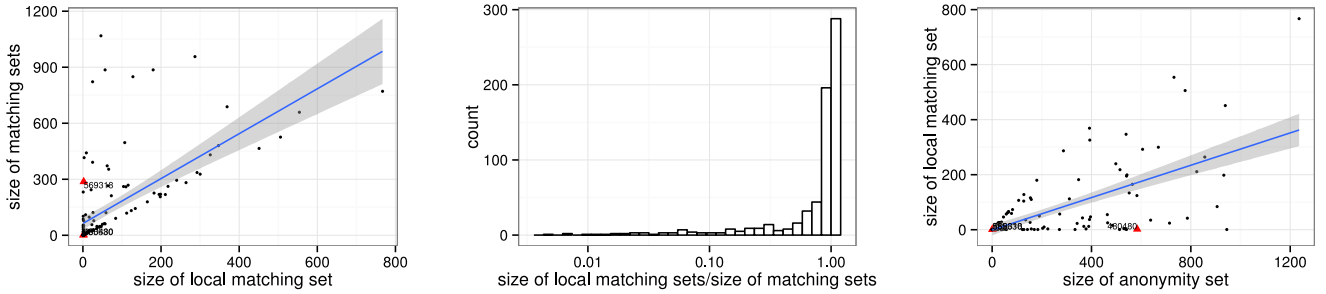


Figure 11: The fraction of the anonymous set sizes to the corresponding matching set sizes $\frac{|\mathcal{A}(d)|}{|\mathcal{M}(d)|}$ where $d = \mathsf{dist}(\mathcal{I}_\mathcal{S}, \mathcal{I}_\mathcal{T})$ over all pairs of subreddits in our data set.

The local matching set is a much better linkability risk measure than the anonymity set because it takes into account the structure of the identity space, i.e., the positioning of identities with respect to each other based on the similarities between them. Figure 12c plots the correlation between the anonymity set size and the local matching set size. They do not correlate perfectly because the identities inside a anonymity set are not distributed uniformly. This makes the anonymity set alone a bad estimate for linkability, because only a fraction of the identities in the anonymity set may be relevant for the matching of two identities. The local matching set, on the other hand, only contains identities that also appear in the matching set $\mathcal{M}$ and thus provide a lower bound for its size.

5.6 Discussion

The insights obtained through the experimental evaluation are twofold: First, the consistent ranking by anonymity set size we derived in Section 3.1, while showing some positive correlation with the matching set size, does not provide a sufficient assessment of the linkability of a user's identities across social media platforms. Second, extending (k, d)-anonymity with information about the matching identity re-



(a) Comparison of the sizes of local matching sets and matching sets.

(b) The conformity of local matching set sizes to the corresponding matching set sizes $\frac{|\mu|}{|\mathcal{M}(d)|}$ over all pairs of subreddits in our data set.

(c) Size of the local matching set compared to the size of the anonymous subset.

Figure 12: Absolute anonymity evaluation.

sults in local matching sets that provide a good approximation of the absolute risk for matching identities to be linked.

Insufficiency of Anonymity By our first insight, we can conclude that only considering the anonymity of an identity within domain is not sufficient to assess the likelihood of linking matching identities across domains. This is contrary to the traditional database setting, where we have a strong relation between anonymity in linkability due to the predefined and restricted number features (i.e. columns in the database) and the required exact equality for anonymity [22].

Such a results was to be expected: the linking process itself utilizes much more information than what is used for determining the anonymity of an identity within a community. As discussed by Goga et al. [12], properties that allow for a successful matching (for instance availability and consistency of identity attributes) depend on both source and target identity. Therefore focusing only on the target identity and its anonymity would not be sufficient to provide a good estimate of linkability.

Absolute Linkability Measure By their definition, anonymity sets, local matching sets and matching sets can only grow by increasing the number of identities within a domain. In practice, this means that, even in very large social media systems with millions of users, it is sufficient to only determine the local matching set size on parts of the whole system. Further increasing the number of considered identities can only increase local matching set and matching set size due to their monotone nature. We therefore only need to gather as much data as is necessary to achieve the linkability estimates that we require.

Defensive Mechanisms From our evaluations, we can also infer directions for potential defensive mechanisms against linkage. In general, users should try to increase the distance between their matching identities since this also increases the matching set size, and therefore decreases the potential linkability.

Furthermore, users should try and avoid exhibiting unique features. We observed in our evaluations for the absolute measure in Section 5.5 that the anonymity set $\mathcal{A}$ of an identity alone is not a good estimate of linkability due to the potentially uneven distribution of identities in $\mathcal{A}$. Such an uneven distribution can be caused by unique attributes that are exhibited by the source and target identity, but not by other identities in the target domain. Goga et al. [12] capture this under the notion discriminability of attributes.

Limitations In our evaluations, we represent identities with a unigram identity model based on the comments users post on Reddit. As discussed in Section 2, users also share other types of content, such as audio, video and text content that can be analyzed in a much more elaborate manner. Including more features of user-content into our analysis will induce different identity models with a (possibly different) corresponding distance measure. We expect our anonymity measures to be applicable to such different settings, since they rely on the relation between anonymity set and (local) matching sets that also hold for other metric spaces than the one considered in this paper.

6. RELATED WORK

We already discussed the primarily relevant, related work in Section 2. Before concluding this paper, we give a short overview of further related work.

Anonymity in Social Networks There are a number of studies which explored the anonymity of nodes in social networks. The focus of theses studies is to investigate how anonymous a given node is in a graph structure [26, 8]. At a high level, the studies show how to transform and apply notions such as k-anonymity to anonymize social graphs by removing or adding edges such that each node in the graph is indistinguishable in a set of k other nodes. Since these studies only consider the social structure of social networks as quasi-identifier, the works are more related to anonymity in traditional databases than anonymity in social media systems.

Matching Identities A number of works propose matching schemes that leverage profile attributes provided by users themselves such as their names, locations or bios [12, 21, 3, 20]. Of particular interest is the study by Goga et al. [12], which shows that it is possible to accurately link 30% of Twitter identities to their matching identities on Facebook. However, it is not possible to exactly pinpoint the matching identity for the remaining 70% of Twitter identities. This insight serves as perfect motivation for our paper: can we build a framework that assesses the individual risk of a user to have his identities matched across sites.

Other studies matched identities by exploiting friends lists or the graph structure of social networks [25, 15, 14, 19]. For example, Narayanan et al. [19] showed the feasibility to de-anonymize the friendship graph of a social network on a large scale using the friendship graph of another social network as auxiliary information.

For geo-location data specifically, Cecaj et al [7] investigate the possibility of matching identities in call detail records to identities in social networks. They characterize the uniqueness of an identity by the number of data points required to uniquely identify an identity and then try to match this uniquely identified identity to its social network profiles using statistical methods similar to the ones proposed in this paper.

Finally, other papers proposed schemes to match identities by exploiting the content generated by users [11, 18, 13]. For example, Mishari et al. [18] show that domain reviews could also be linked across different sites by exploiting the language model of the authors.

Several other works show that even stylometric features of text can be leveraged to identify the author of a given text [4]. Inspired by these works, we also use language models to represent identities. Note that our risk assessment framework can work with any kind of attributes, but for this study we limit ourselves at using language models as attributes.

7. CONCLUSION & FUTURE DIRECTIONS

In this paper, we investigate whether anonymity within a social media system is sufficient to protect against the linking of a user's identities across social media sites. To this end, we presented two novel approaches to estimate the linkability of identities by their anonymity within their communities. The relative measure provides a ranking of identities only based on their intra-domain distances to other identities in the same domain. The absolute linkability measure, on the other hand, uses information about the matching identities of the same user in a different social media domain to provide a better estimate of both identities' linkability.

We empirically evaluate both measures on a data set of user-generated text content collected from Reddit. We show that, on the one hand, the relative measure that relies on anonymity alone is not sufficient for assessing linkability. The absolute measure, on the other hand, provides a meaningful assessment of linkability suitable for application in practice: it does not rely on information about all identities within a social media system, but instead can be evaluated on a subset of all identities.

In addition to the directions discussed in Section 5.6, we consider the following direction important for future work: in practice, social media systems have an ever-changing set of identities that participate, while in this work we consider a static set of identities. Therefore, an efficient method for computing and updating anonymity sets needs to be developed to deal with the dynamically changing nature of social media systems.

8. ACKNOWLEDGEMENTS

This work was supported by the German Federal Ministry of Education and Research (BMBF) through funding for the Center for IT-Security, Privacy and Accountability (CISPA) (FKZ: 16KIS0656), by the German Research Foundation (DFG) via the collaborative research center "Methods and Tools for Understanding and Controlling Privacy" (SFB 1223), project A5, and the European Research Coucil (ERC) Synergy Grant imPACT (no. 610150).

9. REFERENCES

[1] The online social network reddit. http://www.reddit.com. Accessed Nov 2014.

[2] Spokeo. http://www.spokeo.com/.

[3] Alessandro Acquisti, Ralph Gross, and Fred Stutzman. Face recognition and privacy in the age of augmented reality. *Journal of Privacy and Confidentiality*, 6(2):1, 2014.

[4] Sadia Afroz, Michael Brennan, and Rachel Greenstadt. Detecting Hoaxes, Frauds, and Deception in Writing Style Online. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy (S&P)*, pages 461–475, 2012.

[5] Michael Backes, Pascal Berrang, Oana Goga, Krishna Gummadi, and Praveen Manoharan. On Estimating Linkability through Anonymity in Social Media Systems - Supplementary Material. https://infsec.cs.uni-saarland.de/projects/reddit_anonymity/.

[6] Michael Backes, Pascal Berrang, and Praveen Manoharan. From Closed-world Enforcement to Open-world Assessment of Privacy. http://arxiv.org/abs/1502.03346, 2016. eprint arXiv:1502.03346 – cs.CR.

[7] Alket Cecaj, Marco Mamei, and Franco Zambonelli. Re-identification and information fusion between anonymized cdr and social network data. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–14, 2015.

[8] James Cheng, Ada Wai-chee Fu, and Jia Liu. k-isomorphism: Privacy Preserving Network Publication Against Structural Attacks. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data (SIGMOD)*, pages 459–470, 2010.

[9] Cynthia Dwork. Differential Privacy. In *Proceedings of the 33rd International Coloqium on Automata, Languages and Programming (ICALP)*, pages 1–12, 2006.

[10] Dominik M. Endres and Johannes E. Schindelin. A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49(7):1858–1860, 2003.

[11] Oana Goga, Howard Lei, SHK. Parthasarathi, Gerald Friedland, Robin Sommer, and Renata Teixeira. Exploiting innocuous activity for correlating users across sites. In *Proceedings of the 22nd International Conference on World Wide Web (WWW)*, 2013.

[12] Oana Goga, Patrick Loiseau, Robin Sommer, Renata Teixeira, and Krishna P. Gummadi. On the reliability of profile matching across large online social networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2015.

[13] Tereza Iofciu, Peter Fankhauser, Fabian Abel, and Kerstin Bischoff. Identifying users across social tagging systems. In *Proceedings of the 5th*

International Conference on Weblogs and Social Media (ICWSM), 2011.

[14] Nitish Korula and Silvio Lattanzi. An efficient reconciliation algorithm for social networks. *Proceedings of the VLDB Endowment*, 7(5):377–388, 2014.

[15] Sebastian Labitzke, Irina Taranu, and Hannes Hartenstein. What your friends tell others about you: Low cost linkability of social network profiles. In *Proceedings of the 5th International ACM Workshop on Social Network Mining and Analysis (SNA-KDD)*, 2011.

[16] Ninghui Li and Tiancheng Li. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Proceedings of the 23rd International Conference on Data Engineering (ICDE)*, 2007.

[17] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. L-Diversity: Privacy Beyond K-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 2007.

[18] Mishari Al Mishari and Gene Tsudik. Exploring linkability of user reviews. In *Proceedings of the 17th European Symposium on Research in Computer Security (ESORICS)*, 2012.

[19] Arvind Narayanan and Vitaly Shmatikov. De-anonymizing Social Networks. In *Proceedings of the 30th IEEE Symposium on Security and Privacy (S&P)*, pages 173–187, 2009.

[20] Carlton T Northern and Michael L Nelson. An unsupervised approach to discovering and disambiguating social media profiles. In *Proceedings of the 2011 Workshop on Mining Data Semantics (MDS)*, 2011.

[21] Daniele Perito, Claude Castelluccia, Mohamed Ali Kâafar, and Pere Manils. How Unique and Traceable Are Usernames? In *Proceedings of the 10th International Symposium on Privacy Enhancing Technologies (PETS)*, 2011.

[22] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf, 2010. v0.34.

[23] Yanir Seroussi, Ingrid Zukerman, and Fabian Bohnert. Authorship attribution with topic models. *Computational Linguistics*, 40(2):269–310, 2014.

[24] Latanya Sweeney. K-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002.

[25] Gae-won You, Seung-won Hwang, Zaiqing Nie, and Ji-Rong Wen. Socialsearch: enhancing entity search with social network matching. In *Proceedings of the 14th International ACM Conference on Extending Database Technology (EDBT)*, 2011.

[26] Bin Zhou and Jian Pei. The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks. *Knowledge and Information Systems*, 28(1):47–77, 2011.