DeCor: Decorrelating unsubscribed users' data from their identities

Anonymous Authors

1 Introduction

1.1 Motivation

Web application companies face increasing legal requirements to protect users' data. These requirements pressure companies to properly delete and anonymize users' data when a user requests to *unsubscribe* from the service (i.e., revoke access to their personal data). For example, the GDPR requires that any user data remaining after a user unsubscribes is *decorrelated*, i.e., cannot be (directly or indirectly) used to identify the user [4].

In this paper, we propose DeCor, a new approach to managing user identities in web applications. DeCor meets the de-identification requirements in the GDPR, and goes beyond: with DeCor, it is possible for users to switch between a privacy-preserving unsubscribed mode and an identity-revealing subscribed mode at any time. This facilitates important new web service paradigms, such as users granting a time-limited "lease" of data to a service instead of having a permanent service account.

1.2 Goals

DeCor's goal is to provide the following properties:

Decorrelation. Informally, decorrelation should guarantee that it is impossible to distinguish between two records formerly associated with the same unsubscribed user and two records from different unsubscribed users.

Resubscription. Users should be able to easily switch between a privacy-preserving unsubscribed mode and an identity-revealing subscribed mode, without permanently losing their application data.

DeCor must implement these properties while ensuring (1) performance comparable to today's widely-used databases, and (2) easy adoption (requiring little to no modification of application schemas or semantics).

1.3 Threat Model

We make the following assumptions of an attacker whose goal is to break decorrelation:

- An attacker can perform only those queries allowed by the application API: an attacker can access the application only via its public interface. [lyt: Alternatively, an attacker could perform arbitrary queries on some public subset of the application schema (e.g., all tables other than the mapping table, or all tables marked with some compliance policy); arbitrary queries over the entirety of the table are out of scope, unless "private" tables are removed and stored by unsubscribing users.]

- An attacker cannot perform application queries to the past or search web archives: information from prior application snapshots may reveal exactly how data records were decorrelated from unsubscribed users.

- An attacker cannot gain identifying information from arbitrary user-generated content (for example, a reposted screenshot, or text in user posts or comments). Decorrelation seeks to remove identifying information from user-generated data that can be enumerated or follows a specific pattern (e.g. a birthday or email address), and application metadata (e.g., date of postings, database ID columns).

1.4 A General Model for Decorrelation

Application data is structured as tables, each containing a different *data entity*, e.g., a post, user, or vote. Entities can either be externally added via the application API by a client, or internally added by the application (e.g., tags or application metadata). Queries write, read from, and compute over entities. As entities are processed by a query, their path from table to application client can be represented as *flows* through computations specified by the query: each computation can be classified as either a join, an aggregation (e.g., count, min,

max), or a set operation (e.g., union, intersection) between flows.

When a user U unsubscribes, decorrelation must (1) determine which entities may reveal identifying information about U, and (2) prevent queries from leaking this information.

To address (1), the programmer developer specifies identifying tables or table columns via developer annotations on the application schema: tables or table columns can be marked as identity-sensitive. If only tables are marked, all columns are assumed to be sensitive.

To address (2), we model queries as a dataflow computation, and perform actions on sensitive data flowing into computation nodes (joins, aggregations, or set operations) to prevent the computation result from exposing identifying information. Examples of actions include:

- Deleting sensitive entities, or removing sensitive entities from the dataflow. Pros: achieves leave-no-trace. Cons: can disrupt application semantics, and removes any useful information from an unsubscribed user.
- Anonymizing sensitive entity columns: integer or string fields can be randomized, or generated from a distribution of field values (e.g., a phone number).
- Injecting fake entities (noise) or fake associations (changing foreign key values).

[lyt: This seems similar to IFC declassification in a DP/probabilistic sense?]

Note that only certain computations may leak information, and certain actions may need to take place on only one of the flows into the computation instead of both to prevent sensitive information leakage. Here, we list all possible scenarios, and which actions should be taken on which entities flowing into the computation:

- Aggregation/pipelined computations: depend on only one flow of entities. If the entities or entity columns are sensitive, the result is also considered sensitive, and should be acted upon appropriately before reaching the end user.
- JOINs: if one or both of the flows into the joins include sensitive entities or entity columns, all results of the JOIN depend on sensitive entities and should be acted upon before reaching the end user.
- Unions: only the sensitive entities or entity columns must be acted upon in the resulting union
- Intersection: if one of the flows is sensitive, the resulting flow must be acted upon appropriately
- Minus: if either the entities in the subtrahend or the entities in the minuend are sensitive or have sensitive columns, the resulting set of entities must be acted upon. Information can leak about the entities which have been subtracted, or about the entities which remain.

1.5 DeCor: Instantiating the Decorrelation Model

DeCor is one potential architecture for implementing this model. In DeCor, sensitive columns are user ID fields (annotated by the application developer). These user IDs are numerical user keys uid_{key} that are each unique to one user, and map entities to a particular user. Entities containing multiple uid_{key}s are considered shared among the identified users.

DeCor handles flows that contain sensitive entities belonging to unsubscribed users by anonymizing the uid_{key} column: the uid_{key} is replaced by a unique ghost ID (gid_{key}) per entity belonging to the unsubscribed user. By replacing uid_{key} values with gid_{key} values, DeCor ensures that the computations on the dataflow decorrelate these entities with the unsubscribed user: an unsubscribed user's data is split into individual pieces.

While the model implies that actions are taken at runtime (when entities pass through query operators), DeCor implements actions on uid_{key}s by storing the gid_{key}s in the underlying application data tables, saving a mapping of gid_{key}s to uid_{key} in the database, and maintaining materialized views to answer application queries that expose real or ghost identifiers depending on whether a particular user is subscribed. DeCor implements a database shim layer that transparently rewrites application queries to query the materialized view (the "acted-upon" result) rather than the data tables, and which propagates updates appropriately to the materialized view. In essence, the materialized views cache the result of the actions taken after a query operation such as a JOIN.

An alternative action might replace all unsubscribed users' uid_{key}s with one *global placeholder* value, essentially collapsing all unsubscribed users' entities into one pool. However, this erases all user-specific data, making resubscription and subsequent recorrelation of a user's entities with the user's identity difficult. DeCor's ghost identifiers allows users to reactivate their account and undo the decorrelation: uid_{key}s can be linked back to a set of unique gid_{key}s. This gives users the ability to freely unsubscribe to protect their privacy without worrying about losing their accounts. [lyt: Ghosts also make schema changes / changing the location of data records easier to support.]

Furthermore, ghost IDs provide increased decorrelation guarantees: if decorrelation utilizes a global placeholder and entities with uid_{key}s are exposed by application queries, an attacker can determine with 100% certainty which queried entities are unsubscribed entities, namely any entity with uid_{key} equal to the global placeholder. If only one or a few users have unsubscribed, correlating these unsubscribed entities back to one identity may be trivial.

However, if decorrelation instead relies on ghosts, an attacker cannot guarantee that any revealed identifier is a ghost instead of a real (subscribed) user. Instead, the attacker must calculate the probability that an identifier belongs to a ghost

using external knowledge about identifiers (e.g., GIDs may be randomly generated in a identifiable pattern) or other information exposed by the entity. This external, non-uid_{key} information can be either marked sensitive (so DeCor with act upon these entity columns as well) or treated as out of scope.

2 Background and Related Work

2.1 Differentially Private Queries

PINQ, a data analysis platform created by McSherry [2], provides formal differential privacy guarantees for any query allowed by the platform. Data analysts using PINQ can perform transformations (Where, Union, GroupBy, and restricted Join operations) on the underlying data records prior to extracting information via aggregations (e.g., counts, sums, etc.) The aggregation results include added noise to meet the given privacy budget ε, ensuring that analysts only ever receive ε-differentially-private results. PINQ calculates the privacy loss of any given query based on transformations and aggregations to be performed; if a particular query exceeds a predefined privacy budget, PINQ refuses to execute the query.

Differential privacy provides a formal framework that defines the privacy loss a user incurs when the user's data is included in the dataset. However, the setting of differential privacy that of DeCor in several important ways.

First, web applications' utility often derives from visibility into individual data records. PINQ restricts JOIN transformation and exposes information only through differentially-private aggregation mechanisms. While sufficient for many data analyses, this approach severely hinders web application functionality. DeCor supports all application queries, even ones that expose individual records. However, this makes formally defining the privacy guarantees of DeCor more complex than simply applying DP. DP provides formal guarantees for results such as sums or averages because such results are computed using well-defined mathematics, allowing us to neatly capture the total knowledge gained by the adversary. However, in the world of web applications, the knowledge gained via queries that may reveal individual data records cannot be so easily defined: adversaries may learn information from the friends who liked the (ghosted) story, or the time and location the story was posted.

Thus, we cannot, in general, use the DP approach of asking, "by how much do answers to adversary queries change with the presence of a users' decorrelated data?" While we can precisely define which data records an adversary sees with and without a user's decorrelated data, we cannot numerically quantify the knowledge gained by the adversary in the same way as we can quantify the percentage change to a sum or average. Instead, DeCor reduces identifying information in the changes induced by a users' (decorrelated) data, rather than reducing the amount of change itself.

This is demonstrated by how PINQ and DeCor differ in masking entities associated with unique keys. PINQ restricts JOINs to group by unique keys. As an (imperfect) example, selecting a JOIN of stories with users via the UID would result in a record, one per UID, which represents the "group" of all stories associated with a user. A normal join would create a separate record per story, each associated with the user of that UID.

Instead of lumping all matching stories together, DeCor explodes the users' UID into individual unique keys (ghosts), one per story. By PINQ's analysis of stable transformations, this leads to unbounded stability (as the number of different results produced by queries is proportional to the potentially very large number of stories for that user). In PINQ's DP framework, such unbounded stability leads to unbounded privacy loss: one users' data could change aggregation results in "significant" ways unless huge amounts of noise were added.

While DeCor's approach has been shown to be problematic in the past (psuedo/anonymization techniques still allow for reidentification/reconstruction attacks), DeCor provides better anonymization than prior approaches, which decorrelate only coarsely by associating remnants of data with a global placeholder for all deleted users, or do not decorrelate at all (e.g., replacing usernames with a pseudonym, as in the AOL dataset case [**?**]). In addition, DeCor and PINQ's approaches can be complementary: for queries that are more analytic in nature (for example, population statistics used by the web application to measure usage or trends), DeCor can utilize PINQ's technique to provide differentially private guarantees (for a particular snapshot of the dataset). DeCor's decorrelation, in fact, adds more noise than necessary in these cases: a count of unique users who posted about cats will be more imprecise because each post by a decorrelated user is now associated with a unique (ghost) user.

2.2 Deletion Privacy

The right to be forgotten has also been formally defined by Garg et al. [1], where correct deletion corresponds to the notion of leave-no-trace: the state of the data collection system after a user requests to be forgotten should be left (nearly) indistinguishable from that where the user never existed to begin with. While DeCor uses a similar comparison, Garg et al.'s formalization assumes that users operate independently, and that the centralized data collector prevents one user's data from influencing another's.

Other related works:

- Problems with anonymization + reidentification later on (see PINQ for references)
- Deceptive Deletions for protecting withdrawn posts: https://arxiv.org/abs/2005.14113
- "My Friend Wanted to Talk About It and I Didn't": Understanding Perceptions of Dele-

tion Privacy in Social Platforms, user survey https://arxiv.org/pdf/2008.11317.pdf; talk about decoy deletion, prescheduled deletion strategies [3]

- Contextual Integrity
- ML Unlearning
- k-anonymization, pseudonymization

3 Design

4 Implementation

5 Evaluation

6 Discussion

Acknowledgments

Availability

[lyt: USENIX program committees give extra points to submissions that are backed by artifacts that are publicly available. If you made your code or data available, it's worth mentioning this fact in a dedicated section.]

References

[1] Sanjam Garg, Shafi Goldwasser, and Prashant Nalini Vasudevan. Formalizing data deletion in the context of the right to be forgotten. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 373–402. Springer, 2020.

[2] Frank McSherry. Privacy integrated queries. *Communications of the ACM*, 53:89–97, September 2010.

[3] M. Minaei, Mainack Mondal, and A. Kate. "my friend wanted to talk about it and i didn't": Understanding perceptions of deletion privacy in social platforms. *ArXiv*, abs/2008.11317, 2020.

[4] E Parliament. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). *Official Journal of the European Union*, L119(May 2016):1–88, 2016. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC.