Poster: Forgetting the Forgotten, Concealing Content Deletion from Persistent Observers

Title: Lethe: Conceal Content Deletion from Persistent Observers [1]
Authors: M. Minaei, M. Mondal, P. Loiseau, K. Gummadi, and A. Kate
Date: 2019
Venue: Proceedings on Privacy Enhancing Technologies
DOI: https://doi.org/10.2478/popets-2019-0012

Abstract Most social platforms offer mechanisms allowing users to delete their posts, and a significant fraction of users exercise this right to be forgotten. However, ironically, users' attempt to reduce attention to sensitive posts via deletion, in practice, attracts unwanted attention from stalkers specifically to those (deleted) posts. Thus, deletions may leave users more vulnerable to attacks on their privacy in general. Users hoping to make their posts forgotten face a "damned if I do, damned if I don't" dilemma. Many are shifting towards ephemeral social platform like Snapchat, which will deprive us of important user-data archival. In the form of intermittent withdrawals, we present, *Lethe*, a novel solution to this problem of (really) forgetting the forgotten. If the next-generation social platforms are willing to give up the uninterrupted availability of *non-deleted* posts by a very small fraction, *Lethe* provides privacy to the deleted posts over long durations. In presence of *Lethe*, an adversarial observer becomes unsure if some posts are permanently deleted or just temporarily withdrawn by *Lethe*; at the same time, the adversarial observer is overwhelmed by a large number of falsely flagged undeleted posts. To demonstrate the feasibility and performance of *Lethe*, we analyze large-scale real data about users' deletion over Twitter and thoroughly investigate how to choose time duration distributions for alternating between temporary withdrawals and resurrections of non-deleted posts. We find a favorable trade-off between privacy, availability and adversarial overhead in different settings for users exercising their right to delete. We show that, even against an ultimate adversary with an uninterrupted access to the entire platform, *Lethe* offers deletion privacy for up to 3 months from the time of deletion, while maintaining content availability as high as 95% and keeping the adversarial precision to 20%.

References

[1] M Minaei, M Mondal, P Loiseau, K Gummadi, A Kate, *"Lethe: Conceal Content Deletion from Persistent Observers"*, Proceedings on Privacy Enhancing Technologies, 2019.

LETHE: CONCEAL CONTENT DELETION FROM PERSISTENT OBSERVERS

Mohsen Minaei[1], Mainack Mandal[2,3], Patrick Loiseau[2],
Krishna Gummadi[2], Aniket Kate[1]

[1]Purdue University, USA [2]MPI-SWS, Germany [3]University of Chicago, USA

1. DATA ENORMITY & DATA EXPOSURE

People freely talk about their personal life and opinions on online social platforms, causing **long-term exposure** to intended recipients and unintended data scavengers.

2. PRIVACY ISSUES AND DELETIONS

Privacy Attacks.

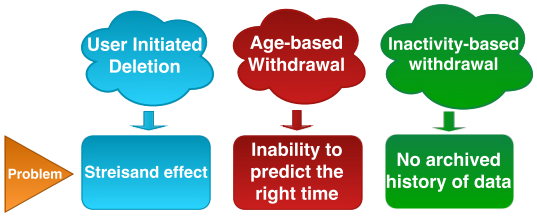
- Harassment and blackmail of victims for their sensitive posts
- Given the enormity of accessible data, large scale mining may not be economical for the adversary

Data Deletions.

- Users request to withdraw their data are honored by several platforms and archives
- *Streisand effect*: an attempt to hide some information has the unintended consequence of bringing particular attention of public to it
- Data deletions make the data scavengers task significantly easy!

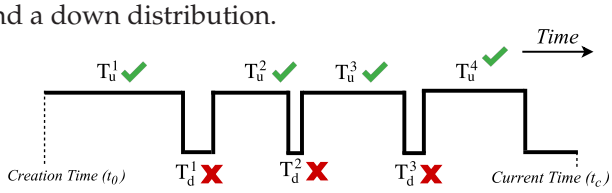
"damned if I do, damned if I don't"

3. STATE OF THE ART AND PROBLEMS



4. LETHE: INTERMITTENT WITHDRAWALS

Our proposal employs an *intermittent withdrawal mechanism* using two time distributions: an up distribution, and a down distribution.



T_u^i is the duration of an up phase(post is visible to all) and T_d^i is the duration of a down phase(post is hidden to all). we toggle between the up and down durations as long as it has not been deleted

5. SECURITY GOALS IN LETHE

Deletion Privacy

- Uncertainty about a post being deleted or just temporarily withdrawn.
- Defined as a likelihood ratio (LR) between a post being in a down duration vs. the post being deleted at a particular time.

Platform availability

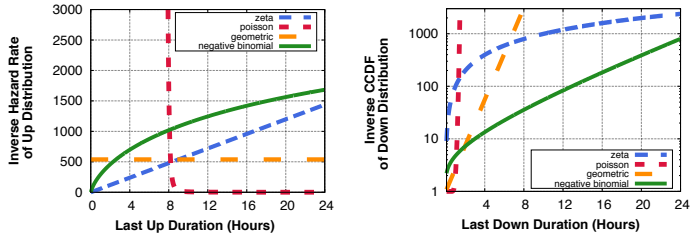
- Average availability of a post within a period.
- Providing privacy guarantees to users while obtaining high availability.

Adversarial Overhead

- # of non-deleted posts falsely flagged as deleted (false-positives) that adversary has to investigate along with the detected actual deleted posts (true-positives).
- Captured by the *precision* measure,

$$\text{precision} = \frac{\text{true-positives}}{\text{true-positives} + \text{false-positives}}$$

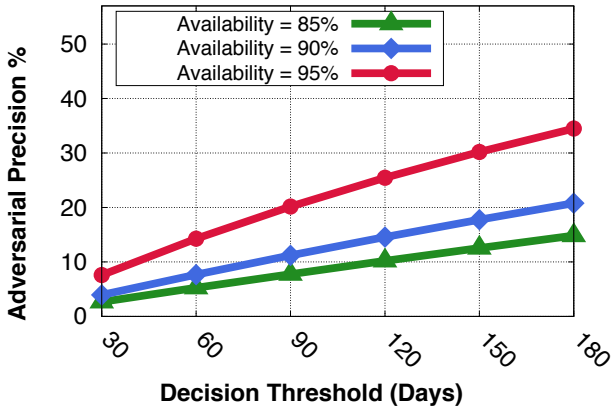
6. DISTRIBUTION SELECTION



Picking **geometric** and **negative binomial** distributions as the up and down distribution, achieves good privacy guarantees, and reasonable availability.

7. EXPERIMENTAL RESULTS

- Evaluated the effectiveness of Lethe with interaction data from Twitter.
- Showing the trade-off between privacy, availability and the adversarial overhead.



Our notion of privacy is simplified to *Decision Threshold*: adversary's wait duration to make a decision about a post being hidden or deleted.