A Queue-based Mechanism for Unlinkability under Batched-timing Attacks

Alexander Goldberg, Giulia Fanti, and Nihar B. Shah

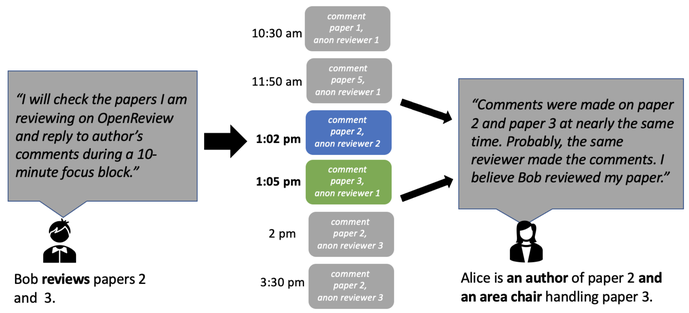
Carnegie Mellon University

Motivated by applications in online peer review forums and cryptocurrency transactions, we consider deanonymization risk arising due to batching—multiple actions taken by a user at nearly the same time. We (1) propose a new formulation of privacy against batching attacks, (2) give an algorithm that introduces delays to the system to preserve privacy, and (3) prove impossibility results under standard differential privacy formulations.

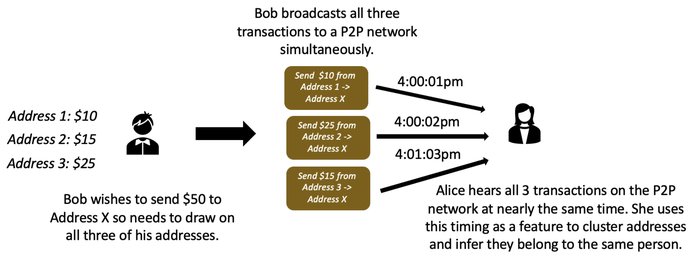
Motivation

- In applications where anonymity is critical, users take public actions under pseudonyms to preserve their privacy.
- When users engage in *batching* — the completion of several similar tasks by the same person at the same time — the simultaneity of their actions may allow an adversary to *link pseudonyms and compromise anonymity.*
- Unlike prior work on linkage attacks in anonymous networks (e.g., [2],[3]) we consider settings where *generating fake data is undesirable.*

Application 1: Inferring the identity of reviewers in anonymous open forums like OpenReview for peer review.

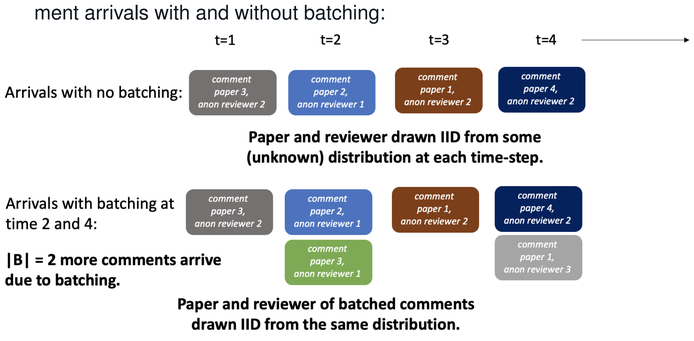


Application 2: Linking cryptocurrency transactions on a blockchain.



Problem Formulation

- Consider **scientific peer review**, where *anonymous reviewers* make *comments* on a set of *papers* on a publicly viewable forum.
- Comments arrive over an *infinite time horizon* in *discrete time*. A **comment arrival sequence** A consists of the sets of comments arriving at each time-step. We consider **two arrival processes** $\mathcal{A}^{(0)}$ and $\mathcal{A}^{(B)}$ that model comment arrivals with and without batching:

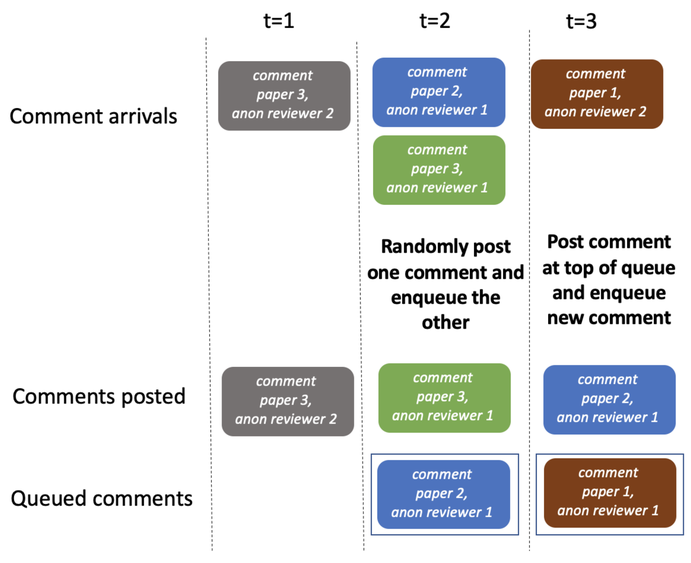


- A **valid comment posting mechanism** respects the following properties:
  (1) (*Delay-Only*) If a comment arrives at time t it must be outputted at time $t' \geq t$.
  (2) (*No Fake Data*) A comment posted at time t arrived at or before time t.
  (3) (*No Withholding Data*) Letting the random delay of a comment be d: $\lim_{m\to\infty} \Pr[d \leq m] = 1$.
- The **utility** of a comment posting mechanism is measured by *worst-case expected delay introduced to a comment.*
- A comment-posting mechanism $\mathcal{M}$ is **ϵ-batching private** with respect to arrival processes $(\mathcal{A}^{(0)}, \mathcal{A}^{(B)})$ if for all time horizons $T \geq 1$, all finite batching multi-sets B, and any output of the mechanism between time 1 and T, C_T:

$$e^{-\epsilon} \leq \frac{\Pr[\mathcal{M}(A) = C_T; A \leftarrow \mathcal{A}^{(B)}]}{\Pr[\mathcal{M}(A) = C_T; A \leftarrow \mathcal{A}^{(0)}]} \leq e^{\epsilon}$$

Algorithm

Algorithm 1 Queue Algorithm
Initialize empty queue $Q = \emptyset$
for t= $1, 2, \ldots$ **do**
  if set of batched comments A arrives **then**
    if $Q \neq \emptyset$ **then**
      Dequeue comment c' from Q and post it.
      Enqueue all comments in A to Q in a random order.
    else
      Choose $c \in A$ uniformly at random to post.
      Enqueue all comments in $A \setminus \{c\}$ to Q in a random order.
      Post comment c immediately.
    end if
  else if a single comment c arrives **then**
    if Q$\neq \emptyset$ **then**
      Dequeue comment c' from Q and post it.
      Enqueue comment c to Q.
    else
      Post comment c.
    end if
  end if
end for



Analysis

Proposition 1. *: When the algorithm is applied to any comment sequence drawn from $\mathcal{A}^{(0)}$ or $\mathcal{A}^{(B)}$:*
(1) (Privacy) The algorithm guarantees perfect batching privacy ($\epsilon = 0$).
(2) (Delay) The worst-case delay of any comment is $|B|$.

Proposition 2. *Any comment-delaying mechanism guaranteeing ϵ-batching privacy with $\epsilon < \infty$ for $\mathcal{A}^{(0)}$ and $\mathcal{A}^{(B)}$ introduces delay of at least $|B|$ to at least one comment when applied to comments arriving according to $\mathcal{A}^{(B)}$.*

Impossibility of Standard Differential Privacy

Why make assumptions on the arrival process of comments instead of using a standard definition of differential privacy?

▷ It is impossible to satisfy the constraints of both a *valid* comment posting mechanism and (ϵ, δ):-differential privacy for various natural notions of "neighbors".

Differential Privacy for Batched Comment Arrivals: A mechanism $\mathcal{M}$ satisfies (ϵ, δ)-differential privacy (DP) [1] if for any finite time horizon T and any set of outputs during this time period S_T,

$$\Pr[\mathcal{M}(A) \in S_T] \leq e^{\epsilon} \Pr[\mathcal{M}(A') \in S_T] + \delta$$

where A and A' are two *"neighboring" comment arrival sequences.*

Definition of "Neighboring" Comment Arrival Sequences	Impossibility Result
Add or remove a batched comment	No *valid* (ϵ, δ)-DP comment posting mechanism for $\epsilon < \infty, \delta < 1$
Move a batched comment to another time-step where it is no longer batched	No *valid* (ϵ, δ)-DP comment posting mechanism for $\epsilon < \infty, \delta < 1$
Move a batched comment by at most g time-steps to another time-step where it is no longer batched	Any *valid* (ϵ, δ)-DP comment posting mechanism delays a comment by $\geq m$ with probability $\geq 1 - 2\delta(m+1)$

Ongoing Work

- *Extend Arrival Model*
- *Privacy Opt-Outs:* allow some users to opt out of privacy-preserving delays
- *Active Adversaries:* protect against adversaries who observe added delay
- *Complementary Formulations of Batching:* a formulation where the set of comments is fixed, but batching varies the timing of a comment
- *Model Interaction of the Mechanism and Arrival Process*

References

[1] C. Dwork et al. "Calibrating Noise to Sensitivity in Private Data Analysis". *Proceedings of Third Conference on Theory of Cryptography*. 2006.

[2] B. Levine et al. "Timing Attacks in Low-Latency Mix Systems". *Financial Cryptography*. 2004.

[3] V. Shmatikov and M. Wang. "Timing Analysis in Low-Latency Mix Networks: Attacks and Defenses". *Computer Security – ESORICS*. 2006.