For Your Eyes Only: Privacy-preserving eye-tracking datasets (Supplementary Material)

Brendan David-John
brendanjohn@ufl.edu
University of Florida
USA

Kevin Butler
butler@ufl.edu
University of Florida
USA

Eakta Jain
ejain@cise.ufl.edu
University of Florida
USA

CCS CONCEPTS

• **Computing methodologies** → *Image processing*; • **Security and privacy** → **Privacy protections**; **Human and societal aspects of security and privacy**; • **Human-centered computing** → *Ubiquitous and mobile computing*; *Human computer interaction (HCI)*.

KEYWORDS

privacy, eye tracking, biometrics, re-identification

ACM Reference Format:
Brendan David-John, Kevin Butler, and Eakta Jain. 2022. For Your Eyes Only: Privacy-preserving eye-tracking datasets (Supplementary Material). In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22), June 8–11, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3517031.3529618

The Supplementary Material provides the formal definition of the Plausible Deniability guarantee achieved by the Marginals Generative Model in our work, as well as a proof of sufficient conditions that establish the guarantee.

1 PLAUSIBLE DENIABILITY PRIVACY CRITERION

Plausible Deniability has two privacy parameters: k and γ [Bindschaedler et al. 2017]. The privacy criterion and assumptions are as follows. Let $\mathbf{M}$ be a probabilistic generative model that takes as input a data record d and generates synthetic records y with probability equal to $Pr\{y = \mathbf{M}(d)\}$.

Definition 1. *Plausible Deniability (PD)*

For any dataset D where $|D| \geq k$, and any record y generated by a probabilistic generative model $\mathbf{M}$ such that $y = \mathbf{M}(d_1)$ for $d_1 \in D$, we state that y is releasable with (k,γ)-PD if there exist at least $k-1$ unique records $d_2, ..., d_k \in D \setminus \{d_1\}$, s.t.

$$\gamma^{-1} \leq \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} \leq \gamma$$

where $k \geq 1$ is an integer and $\gamma \geq 1$ is a real number.

The level of privacy is controlled by parameters k and γ. Large values of k, and values of γ that are closer to one imply higher privacy. In practice, PD ensures that at least $k-1$ plausible seeds, i.e., inputs, to the model $\mathbf{M}$ could have plausibly produced the output synthetic record y. The parameter γ bounds how close together the probabilities are to determine whether they were plausible seeds. Privacy-preserving datasets are generated by only releasing synthetic records y if they pass the PD privacy test.

PD Privacy Test: For each synthetic candidate $y = \mathbf{M}(d)$:

(1) Let i ≥ 0 be the only integer that fits the inequality $\gamma^{-i-1} < Pr\{y = \mathbf{M}(d)\} \leq \gamma^{-i}$
(2) Let k' be the count of records $d_a \in D$ such that $\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_a)\} \leq \gamma^{-i}$
(3) If $k' \geq k$ test result is *pass* and y can be released, else test result is *fail*

Step 1 is formulated as there is only one integer that satisfies the inequality when $\gamma \geq 1$, as the range of values covered by the set $(\gamma^{-i-1}, \gamma^{-i}]$ represent disjoint sections of the real number line for different integer values of i. Therefore, $Pr\{y = \mathbf{M}(d)\}$ can only fall within one such range. Step 2 works because the condition that both $Pr\{y = \mathbf{M}(d)\}$ and $Pr\{y = \mathbf{M}(d_a)\}$ fall within the range of $(\gamma^{-i-1}, \gamma^{-i}]$ is sufficient to satisfy the inequality for (k,γ)-PD.

2 PROOF OF SUFFICIENT CONDITION FOR PLAUSIBLE DENIABILITY

Theorem. For real $\gamma \geq 1$, if

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i} \text{ and } \gamma^{-i-1} < Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$$

are true for the only integer $i > 0$, then

$$\gamma^{-1} \leq \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma.$$

Proof. Assume that

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i} \text{ and } \gamma^{-i-1} < Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$$

for the only integer $i > 0$. Starting with

$$\gamma^{-i-1} < Pr\{y = \mathbf{M}(d_i)\} \leq \gamma^{-i},$$

divide all terms by $Pr\{y = \mathbf{M}(d_j)\}$ to get

$$\frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} \leq \frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}}.$$

Because $Pr\{y = \mathbf{M}(d_j)\} \leq \gamma^{-i}$, we have that

$$\frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} \geq \frac{\gamma^{-i-1}}{\gamma^{-i}} = \gamma^{-1}.$$

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ETRA '22, June 8–11, 2022, Seattle, WA, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9252-5/22/06…$15.00
https://doi.org/10.1145/3517031.3529618

Because $Pr\{y = \mathbf{M}(d_j)\} > \gamma^{-i-1}$, we have that

$$\frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{\gamma^{-i}}{\gamma^{-i-1}} = \gamma$$

therefore,

$$\gamma^{-1} \le \frac{\gamma^{-i-1}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \frac{\gamma^{-i}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma$$

which satisfies

$$\gamma^{-1} \le \frac{Pr\{y = \mathbf{M}(d_i)\}}{Pr\{y = \mathbf{M}(d_j)\}} < \gamma. \qquad \square$$

REFERENCES

Vincent Bindschaedler, Reza Shokri, and Carl A Gunter. 2017. Plausible Deniability for Privacy-Preserving Data Synthesis. *Proceedings of the VLDB Endowment* 10, 5 (2017).