Alma Mater Studiorum · Università di Bologna

Scuola di Ingegneria e Architettura
Dipartimento di Informatica · Scienza e Ingegneria · DISI
Corso di Laurea Magistrale in Ingegneria Informatica

DIFFERENTIAL PRIVACY

Relatore:
Prof. Michele Colajanni

Presentata da:
Riccardo Barbieri

Correlatore:
Silvio Russo

Anno Accademico 2023/2024

Elenco delle figure

Indice

Introduzione

—*Ancora un placeholder, da modificare*— La privacy dei dati è diventata una preoccupazione sempre più importante nell'era digitale. Con la crescente raccolta e l'analisi di dati personali, è fondamentale garantire che le informazioni individuali siano protette da accessi non autorizzati e da usi impropri. La privacy differenziale (DP) è una rigorosa definizione matematica di privacy che offre forti garanzie di riservatezza anche quando i dati vengono utilizzati per l'analisi.

1 | Stato dell'arte

1.1 Introduzione alla Differential Privacy

La privacy differenziale (DP) è un framework per la progettazione di algoritmi utilizzati per la creazione di distribuzioni di dati aggregati su dataset; gli algoritmi che rispettano la definizione di DP sono in grado di limitare l'impatto che la partecipazione di un singolo individuo ha sui risultati dell'analisi, rendendo *quasi* impossibile confermare o negare la presenza di un soggetto all'interno del dataset utilizzato per l'aggregazione di dati.

Questo risultato è ottenuto grazie all'aggiunta di rumore casuale campionato da una distribuzione di probabilità opportuna ai risultati dell'elaborazione dei dati.

Lo scopo principale degli algoritmi DP è quello di poter garantire ai partecipanti di un dataset che la privacy dei dati forniti non potrà essere violata a seguito di analisi effettuate sul dataset, indipendentemente dalla disponibilità di altre fonti di informazioni su uno specifico individuo [1, p. 5].

1.1.1 Definizione

La definizione seguente costituisce un formalismo più ristretto rispetto alla definizione completa, quest'ultima include un secondo parametro δ, la cui funzione e utilità verrà discussa in seguito.

Una trasformazione $\mathcal{M}$ è considerata ϵ-differenzialmente privata se, per ogni dataset D_1 e D_2 che differiscono solo per un individuo, vale la seguente disequazione [2]:

$$\Pr[\mathcal{M}(D_1) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D_2) \in S] \tag{1.1}$$

La definizione è costituita da tre termini:

- $\Pr[\mathcal{M}(D_1) \in S]$
- $\Pr[\mathcal{M}(D_2) \in S]$
- e^{ϵ}

Bibliografia

[1] Cynthia Dwork e Aaron Roth. «The Algorithmic Foundations of Differential Privacy». In: *Foundations and Trends® in Theoretical Computer Science* 9.3–4 (2014), pp. 211–407. ISSN: 1551-305X. DOI: 10.1561/0400000042. URL: http://dx.doi.org/10.1561/0400000042.

[2] Cynthia Dwork et al. «Calibrating Noise to Sensitivity in Private Data Analysis». In: *Theory of Cryptography*. A cura di Shai Halevi e Tal Rabin. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. ISBN: 978-3-540-32732-5.