Differentially Private Selection using Smooth Sensitivity

Iago Chaves, Javam Machado

Laboratório de Sistemas e Banco de Dados (LSBD)
Departamento de Computação / UFC – Fortaleza – CE – Brazil

{iago.chaves, javam.machado}@lsbd.ufc.br

Abstract. *Differentially private (DP) selection mechanisms identify the top-scoring element from a finite set while ensuring strong privacy guarantees. However, most existing methods rely on global sensitivity, often introducing excessive noise and harming utility. We propose the Smooth Noisy Max (SNM) algorithm, which leverages smooth sensitivity to provide tighter error bounds and improved accuracy under mild conditions. Experiments demonstrate SNM's superiority over state-of-the-art DP selection methods in percentile selection, greedy decision trees, and random forest applications.*

Resumo. *Mecanismos de seleção diferencialmente privados (DP) obtem o elemento de maior qualidade em um conjunto finito, garantindo forte privacidade. No entanto, a maioria dos métodos existentes usa sensibilidade global, frequentemente adicionando ruído excessivo e prejudicando a utilidade dos resultados. Propomos o algoritmo Smooth Noisy Max (SNM), que emprega sensibilidade suave para obter limites de erro mais precisos e maior acurácia sob condições moderadas. Experimentos mostram que o SNM supera os melhores métodos de seleção com DP em três aplicações: seleção de percentis, árvores de decisão gananciosas e florestas aleatórias.*

1. Dados da Defesa da Tese e Pontos de Destaque

Categoria	Programa	Orientador	Data da defesa
Doutorado	Mestrado e Doutorado em Ciência da Computação da Universidade Federal do Ceará	Javam de Castro Machado	28 de agosto de 2024

Membro da Banca	Filiação
Victor Aguiar Evangelista	Universidade Federal do Ceará
César Lincoln Cavalcante Mattos	Universidade Federal do Ceará
Daniel Cardoso Moraes de Oliveira	Universidade Federal Fluminense
Diego Mesquita	Fundação Getulio Vargas

Pontos de Destaque:

- Avanço em IA Responsável: Privacidade e Fairness.
- Algoritmo Smooth Noisy Max (SNM) inova aplicando sensibilidade suave, reduzindo erros.
- SNM garante privacidade diferencial aproximada, pioneiro nessa capacidade.
- Melhorias em seleção diferencialmente privada.
- 3 novos algoritmos de inteligência artificial que superam estado-da-arte.

Nota e Menção honrosa: A instituição não atribui nota e menção honrosa.

2. Contexto e Problema

A gestão responsável de dados é crucial para o desenvolvimento ético de sistemas de IA, pois os algoritmos refletem diretamente a qualidade e o tratamento dos dados utilizados. A falta de governança adequada pode perpetuar vieses discriminatórios (*fairness*) e violações de privacidade, exigindo técnicas robustas de proteção. A adoção de **IA Responsável** no gerenciamento de dados vai além da conformidade: é um imperativo para construir confiança, distribuir benefícios de forma justa e evitar impactos sociais negativos, garantindo inovação tecnológica equitativa e benéfica para a sociedade. A IA Responsável exige técnicas que equilibrem utilidade e privacidade, como a seleção diferencialmente privada. Essa abordagem adiciona aleatoriedade controlando a operações de seleção (ex.: qual a cor de cabelo mais comum no Brasil?), preservando dados individuais, garantindo os requisitos éticos de privacidade. Seleção diferencialmente privada é um mecanismo que protege dados sensíveis ao introduzir aleatoriedade matematicamente calibrada em processos de seleção (consultas cuja resposta é discreta), garantindo que a saída não revele informações individuais específicas.

3. Objetivo

Esta tese dedicou-se ao avanço dos aspectos éticos em Inteligência Artificial, com contribuições direcionadas a algoritmos de aprendizado de máquina, modelo de *fairness* e técnicas de privacidade de dados. O foco central da tese reside na privacidade de dados, e portanto este documento focará nos resultados alcançados no campo da privacidade. Os estudos de privacidade culminaram na proposição de um novo algoritmo de seleção com privacidade diferencial, denominado Smooth Noisy Max (SNM). Este algoritmo inovador emprega a sensibilidade suave para a adição de ruído calibrado. O trabalho aborda desafios e lacunas cruciais na área ao: (1) estender o conceito de sensibilidade suave, originalmente concebido para dados numéricos, ao complexo cenário de seleção de dados; (2) desenvolver um algoritmo que aplica esta noção estendida de sensibilidade suave; e (3) prover rigorosas garantias teóricas de **privacidade** e **utilidade**. Demonstra-se que o SNM não apenas alcança maior precisão em comparação com alternativas baseadas em sensibilidade global, mas também, sob leves condições, garante um desempenho nunca inferior ao de seus concorrentes.

4. A solução e o avanço no estado-da-arte

Este trabalho introduz o Smooth Noisy Max (SNM), primeiro mecanismo de seleção privada que utiliza sensibilidade suave para calibrar ruído de forma adaptativa, garantindo (ε, δ)-DP e superando métodos tradicionais (Mecanismo Exponencial, Report-Noisy-Max, Permute-and-Flip) que dependem de sensibilidade global (gerando ruído excessivo), ou Local Dampening (limitado por instabilidade e complexidade). Diferente dessas abordagens, o SNM adiciona aleatoriedade proporcional à sensibilidade suave ao invés da global. Alem disso, sob condições amenas o SNM *nunca é pior* que os concorrentes, além de ser o primeiro a suportar (ε, δ)-DP, equilibrando precisão e garantias de privacidade de forma inédita.

5. Avaliação

Este trabalho realizou uma avaliação abrangente do SNM, comparando-o com os principais métodos de seleção privada (Mecanismo Exponencial, Report-Noisy-Max, Permute-and-Flip, Local Dampening) em três aplicações distintas: (1) seleção de percentil, (2)

árvores de decisão gulosa e (3) florestas aleatórias. Os resultados demonstraram a superioridade do SNM, que: (i) Na seleção de percentil (datasets PATENT, HEPTH, INCOME), alcançou erros similares com 85% menos orçamento de privacidade; (ii) Em árvores de decisão (datasets Adult, NLTCS, ACS), obteve maior acurácia na maioria dos cenários; (iii) Em florestas aleatórias (Mushroom, Wine, Compas, etc.), atingiu acurácia comparável à versão não privada mesmo com orçamentos reduzidos, superando consistentemente outros métodos para diversos valores de ε.

6. Contribuições

Sensibilidade local em seleção diferencialmente privada A tese analisou mecanismos de seleção privada com sensibilidades dependentes de dados, destacando limitações no mecanismo de amortecimento local de Farias et al. [2023]. A tese demonstrou que a sensibilidade suave é incompatível com o mecanismo exponencial, contradizendo trabalhos anteriores Fletcher and Islam [2017], que carecem de uma definição formal para seleção privada.

Sensibilidade suave para seleção privada A tese propõe uma nova definição de sensibilidade suave para seleção privada. A formulação adapta a sensibilidade dependente do banco de dados real ao contexto de seleção privada.

Novo mecanismo diferencialmente privado A tese apresenta o Smooth Noisy Max (SNM), um novo mecanismo de seleção diferencialmente privado, inspirado no report-noisy-max. Destaques incluem simplicidade, facilidade de implementação e maior precisão ao utilizar sensibilidade suave, reduzindo a magnitude do ruído. SNM garante privacidade diferencial pura e aproximada. O algoritmo adiciona ruído, proporcional à sensibilidade suave, aos escores de utilidade. Garantias teóricas mostram que SNM nunca é pior que os adversários sob condições amenas. Experimentos empíricos validam seu desempenho frente aos métodos existentes.

Seleção de percentil diferencialmente privado Este trabalho avalia empiricamente o SNM para seleção de percentis com privacidade diferencial, comparando-o com os métodos existentes. A análise de sensibilidade guia a calibração de ruído, com testes em datasets reais. Os resultados demonstram que o SNM reduz o orçamento de privacidade em até 85% sem perda de utilidade, destacando-se em conjunto de dados com repetição de valores. Conclui-se que o SNM supera os métodos atuais em eficiência e acurácia.

Árvores de decisão gulosas diferencialmente privadas Este estudo adapta o algoritmo ID3 de árvores de decisão para privacidade diferencial utilizando SNM, comparando-o com métodos concorrentes. É realizado uma análise de sensibilidade que guia a calibração de ruído. Testes em datasets reais mostraram que as variantes do SNM atingiram maior acurácia, com ganhos de até 8,58% em árvores mais profundas e orçamentos de privacidade restritos. Os resultados comprovam a vantagem da sensibilidade suave no equilíbrio entre privacidade e utilidade.

Florestas aleatórias diferencialmente privadas Este trabalho propõe um algoritmo de floresta aleatória com privacidade diferencial utilizando SNM como mecanismo de seleção. Os principais aspectos incluem: (1) análise de sensibilidade para votação majoritária em folhas e sensibilidade suave baseada na diferença entre frequências de classes; e (2) avaliação abrangente em seis datasets, comparando SNM com os mecanismos padrões através de métricas de acurácia sob diferentes níveis de privacidade.

Produção científica

Os artigos listados nesta seção surgem diretamente da tese em questão.

Artigo	Ano	Qualis	h5-index	Citações
Chaves, E. Farias, Perez, Mesquita, and Machado, " Differentially Private Selection Using Smooth Sensitivity ", 2025 IEEE Symposium on Security and Privacy (SP) , 2025	2025	A1	112	0
Silva, **Chaves**, and Machado, "LAGOON: Achieving bounded individual fairness through classification frequency equalization", J. Braz. Comput. Soc., 2024	2024	A2	35	0
Cabral, Farias, Sena, **Chaves**, Pordeus, Santiago, Sá, Machado, and Madeiro, "An Active Learning Approach for Detecting Customer Induced Damages in Motherboards with Deep Neural Networks", Learning & Nonlinear Models, 2023	2023	B2	11	2
Alves, de Farias, **Chaves**, Chao, Madeiro, Gomes, and Machado, "Detecting Customer Induced Damages in Motherboards with Deep Neural Networks", International Joint Conference on Neural Networks, IJCNN 2022, Padua, Italy, July 18-23, 2022, 2022	2022	A1	64	7
Sena, Praciano, **Chaves**, Brito, Neto, Monteiro, and Machado, "AUDIO-MC: A General Framework for Multi-context Audio Classification", Proceedings of the 24th International Conference on Enterprise Information Systems, ICEIS, 2022	2022	A3	21	1
Chaves, Martins, Praciano, Brito, Monteiro, and Machado, "BPA: A Multilingual Sentiment Analysis Approach based on BiLSTM", Proceedings of the 24th International Conference on Enterprise Information Systems, ICEIS 2022, Online Streaming, April 25-27, 2022, Volume 1, 2022	2022	A3	21	1
Lima, Pereira, **Chaves**, Machado, and Gomes, "Predicting the Health Degree of Hard Disk Drives With Asymmetric and Ordinal Deep Neural Models", IEEE Trans. Computers, 2021	2021	A1	141	10
M. Silva, **C. Chaves**, and C. Machado, "Private Reverse Top-k Algorithms Applied on Public Data of COVID-19 in the State of Ceará", Journal of Information and Data Management, 2021	2021	B1	6	1
Pereira, **Chaves**, Gomes, and Machado, "Using Autoencoders for Anomaly Detection in Hard Disk Drives", 2020 International Joint Conference on Neural Networks, IJCNN 2020, Glasgow, United Kingdom, July 19-24, 2020, 2020	2020	A1	64	6
de Lourdes Maia Silva, **Castro Chaves**, and de Castro Machado, "Aplicação de Top-k Reverso com Privacidade sobre os Dados Públicos de COVID-19 no Estado do Ceará", Proceedings of the 35th Brazilian Symposium on Databases, SBBD 2020, Online, September 28 - October 1, 2020, 2020	2020	A4	7	1
Chaves and Machado, "Differentially Private Group-by Data Releasing Algorithm", Proceedings of the 34th Brazilian Symposium on Databases, SBBD 2019, Fortaleza, CE, Brazil, October 7-10, 2019, 2019	2019	A4	7	0

Referências

Danilo Alves, Victor A. E. de Farias, Iago C. **Chaves**, Richard Chao, João Paulo Madeiro, João Paulo Pordeus Gomes, and Javam C. Machado. Detecting customer induced damages in motherboards with deep neural networks. In *International Joint Conference on Neural Networks, IJCNN 2022, Padua, Italy, July 18-23, 2022*, pages 1–8. IEEE, 2022. doi: 10.1109/IJCNN55064.2022.9892047. URL https://doi.org/10.1109/IJCNN55064.2022.9892047.

L. Cabral, V. Farias, L. Sena, I. **Chaves**, J. P. Pordeus, J. P. Santiago, D. Sá, J. Machado, and J. P. Madeiro. An active learning approach for detecting customer induced damages in motherboards with deep neural networks. *Learning & Nonlinear Models*, 21(2):29–42, 2023. doi: 10.21528/lnlm-vol21-no2-art3.

Victor A. E. de Farias, Felipe T. Brito, Cheryl J. Flynn, Javam C. Machado, Subhabrata Majumdar, and Divesh Srivastava. Local dampening: differential privacy for non-numeric queries via local sensitivity. *VLDB J.*, 32(6):1191–1214, 2023. doi: 10.1007/S00778-022-00774-W. URL https://doi.org/10.1007/s00778-022-00774-w.

Maria de Lourdes Maia Silva, Iago **Castro Chaves**, and Javam de Castro Machado. Aplicação de top-k reverso com privacidade sobre os dados públicos de COVID-19 no estado do ceará. In *Proceedings of the 35th Brazilian Symposium on Databases, SBBD 2020, Online, September 28 - October 1, 2020*, pages 193–198. SBC, 2020. doi: 10.5753/SBBD.2020.13640. URL https://doi.org/10.5753/sbbd.2020.13640.

Sam Fletcher and Md Zahidul Islam. Differentially private random decision forests using smooth sensitivity. *Expert Syst. Appl.*, 78:16–31, 2017. doi: 10.1016/J.ESWA.2017.01.034. URL https://doi.org/10.1016/j.eswa.2017.01.034.

Fernando Dione S. Lima, Francisco Lucas Falcao Pereira, Iago C. **Chaves**, Javam C. Machado, and João Paulo Pordeus Gomes. Predicting the health degree of hard disk drives with asymmetric and ordinal deep neural models. *IEEE Trans. Computers*, 70 (2):188–198, 2021. doi: 10.1109/TC.2020.2987018. URL https://doi.org/10.1109/TC.2020.2987018.

Maria de Lourdes M. Silva, Iago **C. Chaves**, and Javam C. Machado. Private reverse top-k algorithms applied on public data of covid-19 in the state of ceará. *Journal of Information and Data Management*, 12(5), Nov. 2021. doi: 10.5753/jidm.2021.1941. URL https://journals-sol.sbc.org.br/index.php/jidm/article/view/1941.

Francisco Lucas Falcao Pereira, Iago Castro **Chaves**, João Paulo Pordeus Gomes, and Javam C. Machado. Using autoencoders for anomaly detection in hard disk drives. In *2020 International Joint Conference on Neural Networks, IJCNN 2020, Glasgow, United Kingdom, July 19-24, 2020*, pages 1–7. IEEE, 2020. doi: 10.1109/IJCNN48605.2020.9206689. URL https://doi.org/10.1109/IJCNN48605.2020.9206689.

Lucas B. Sena, Francisco D. B. S. Praciano, Iago C. **Chaves**, Felipe T. Brito, Eduardo Rodrigues Duarte Neto, José Maria Monteiro, and Javam C. Machado. AUDIO-MC: A general framework for multi-context audio classification. In Joaquim Filipe, Michal Smialek, Alexander Brodsky, and Slimane Hammoudi, editors, *Proceedings of the 24th International Conference on Enterprise Information Systems, ICEIS*, pages 374–383.

SCITEPRESS, 2022. doi: 10.5220/0011071500003179. URL https://doi.org/10.5220/0011071500003179.

Maria Silva, Iago C. **Chaves**, and Javam C. Machado. LAGOON: achieving bounded individual fairness through classification frequency equalization. *J. Braz. Comput. Soc.*, 30(1):238–251, 2024. doi: 10.5753/JBCS.2024.3468. URL https://doi.org/10.5753/jbcs.2024.3468.

Iago C. **Chaves** and Javam C. Machado. Differentially private group-by data releasing algorithm. In *Proceedings of the 34th Brazilian Symposium on Databases, SBBD 2019, Fortaleza, CE, Brazil, October 7-10, 2019*, pages 271–276. SBC, 2019. doi: 10.5753/SBBD.2019.8835. URL https://doi.org/10.5753/sbbd.2019.8835.

Iago C. **Chaves**, Antônio Diogo Forte Martins, Francisco D. B. S. Praciano, Felipe T. Brito, José Maria Monteiro, and Javam C. Machado. BPA: A multilingual sentiment analysis approach based on bilstm. In Joaquim Filipe, Michal Smialek, Alexander Brodsky, and Slimane Hammoudi, editors, *Proceedings of the 24th International Conference on Enterprise Information Systems, ICEIS 2022, Online Streaming, April 25-27, 2022, Volume 1*, pages 553–560. SCITEPRESS, 2022. doi: 10.5220/0011071400003179. URL https://doi.org/10.5220/0011071400003179.

Iago C. **Chaves**, Victor A. E. Farias, Amanda Perez, Diego Mesquita, and Javam C. Machado. Differentially Private Selection Using Smooth Sensitivity . In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 3969–3987, Los Alamitos, CA, USA, May 2025. IEEE Computer Society. doi: 10.1109/SP61157.2025.00216. URL https://doi.ieeecomputersociety.org/10.1109/SP61157.2025.00216.