Riscos de Privacidade em Dados de Saúde: Investigando Inferência de Atributos Sensíveis de Cidadãos no DATASUS

Ramon G. Gonze,[1,2] **Igor W. Lemes,**[1]
Jussara M. Almeida,[1] **Marcos A. Gonçalves,**[1] **Mário S. Alvim**[1]

[1] Universidade Federal de Minas Gerais (UFMG), Brazil

[2]École Polytechnique de Paris, France

{ramon.gonze, igorlemes, jussara, mgoncalves, msalvim} @dcc.ufmg.br

Abstract. *Statistical dissemination of health data is crucial for the formulation and monitoring of public policies and scientific research, but it presents important challenges regarding the privacy of data subjects. In this work, we formally and experimentally evaluate the risks of inferring sensitive attributes in the DATASUS outpatient procedure dataset, which contains microdata since 1994 to the present day on millions of citizens. We identified serious privacy risks – for example, in some cases it is possible to identify sensitive attributes with an accuracy higher than 90% in almost 30% of the records in the database. These results led to the question of whether the platform is compliant with the Lei Geral de Proteção de Dados (LGPD).*

Resumo. *A divulgação estatística de dados de saúde é crucial para a formulação e o acompanhamento de políticas públicas e a pesquisa científica, mas apresenta desafios significativos quanto à privacidade dos titulares dos dados. Neste trabalho avaliamos formal e experimentalmente riscos de inferência de atributos sensíveis na base de dados de procedimentos ambulatoriais do DATASUS, que contém microdados desde 1994 até os dias atuais sobre milhões de cidadãos. Identificamos sérios riscos à privacidade – por exemplo, em alguns casos é possível inferir atributos sensíveis com uma acurácia superior a 90% em quase 30% dos registros da base. Tais resultados motivam o questionamento sobre se a plataforma encontra-se em conformidade com a Lei Geral de Proteção de Dados (LGPD).*

1. Introdução

A transparência na divulgação de dados governamentais é essencial para a elaboração e o acompanhamento de políticas públicas, a pesquisa científica, o jornalismo, dentre outros objetivos benéficos à sociedade. No Brasil, há uma tradição de se buscar a transparência na divulgação de dados. Em particular, a Lei de Acesso à Informação (LAI) determina ao poder público o dever de garantir amplo acesso à informação, particularmente àquela considerada de interesse coletivo ou geral, a qual deve ser disponibilizada via Internet independentemente de requerimento.

O princípio da transparência, entretanto, deve ser equilibrado com outro valor social: a privacidade dos cidadãos. Em 2014, a Assembleia Geral da Organização das Nações Unidas (ONU) adotou a Resolução 68/261, que estabelece um conjunto de

princípios fundamentais para estatísticas oficiais [Organização das Nações Unidas 2014]. Esta resolução determina que os dados individuais coletados pelos órgãos responsáveis por elaboração de estatísticas, sejam eles referentes a pessoas físicas ou jurídicas, devem ser estritamente confidenciais e utilizados exclusivamente para fins estatísticos. De fato, no mundo todo há esforços legislativos e regulatórios para a proteção da privacidade dos cidadãos, como o *General Data Protection Regulation* (GDPR) europeu [EU 2016], o *Confidential Information Protection and Statistical Efficiency Act* (CIPSEA) dos Estados Unidos [Government of the United States of America 2002], e a revisão australiana de seu *Privacy Act* [Government of Australia 1988]. No Brasil, a Lei Geral de Proteção de Dados Pessoais (LGPD) está plenamente em vigor desde 2021 e prevê sanções para o poder público quando a privacidade dos cidadãos não for protegida em divulgações estatísticas de dados. Neste contexto, impõe-se aos órgãos públicos que coletam, mantêm e divulgam dados estatísticos o desafio de conciliar os requisitos de transparência com os requisitos de proteção de privacidade impostos pela atual legislação brasileira.

Uma maneira formal de se medir a privacidade em um conjunto de dados é através da análise do sucesso de potenciais ataques realistas [Matthews and Harel 2011]. A literatura comumente identifica três tipos principais destes ataques [Alvim et al. 2022]: (i) *pertencimento (membership)*, em que um adversário tenta inferir se os dados de um indivíduo alvo encontram-se presentes ou não em um conjunto de dados; (ii) *reidentificação*, em que um adversário tenta inferir qual indivíduo é o titular de um registro na base de dados; e (iii) *inferência de atributo sensível*, em que o adversário tenta inferir o atributo sensível de um indivíduo alvo, independentemente de se o indivíduo foi reidentificado ou não.

Uma prática comum em órgãos governamentais brasileiros é tentar proteger a privacidade na divulgação de microdados empregando as técnicas de *desidentificação*, em que se removem possíveis identificadores individuais óbvios dos registros (como nome, CPF, RG) e/ou de *pseudonimização*, em que tais identificadores individuais óbvios são substituídos por um código único de identificação artificialmente criado. Entretanto, é sabido na literatura técnica sobre privacidade que, mesmo quando se removem ou se criptografam identificadores explícitos (como nomes, endereços e números de telefone) dos titulares dos microdados em uma divulgação, outros dados distintos, chamados de *quase-identificadores* (QIDs), podem se combinar de maneira inadequada e ser, assim, vinculados a informações publicamente disponíveis para reidentificar os indivíduos.

De fato, Sweeney demonstrou que 87% da população dos EUA poderia ser unicamente reidentificada nos dados do censo demográfico americano de 1990 usando-se como QIDs apenas o sexo, a data de nascimento e o código postal (*zipcode*) dos indivíduos [Sweeney 2000]. No contexto brasileiro, um estudo recente mostrou que a forma de divulgação dos microdados dos Censos Educacionais da Educação Básica e da Educação Superior no Brasil utilizada até 2019 (que empregava exatamente as técnicas de desidentificação e de pseudonimização para proteção de privacidade) permitia a reidentificação e a inferência de atributos sensíveis de uma parcela significativa dos estudantes presentes nas bases [Alvim et al. 2022]. Esta análise mostrou que mais de 60 milhões de cidadãos brasileiros nos Censos corriam riscos elevados de privacidade, o que levou o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), [1]

[1] `https://www.gov.br/inep/pt-br`

curador dos dados, a reavaliar sua forma de publicação para atender a LGPD. [2] [,3]

Neste trabalho voltamos nossa atenção para a plataforma DATASUS,[4] mantida pelo Ministério da Saúde do Brasil, que disponibiliza dados de saúde pública do Sistema Único de Saúde (SUS) em larga escala através da Internet. Entre outras informações, a plataforma permite o acesso a dados sobre hospitalizações, procedimentos ambulatoriais e estatísticas vitais de cidadãos brasileiros. Embora esses conjuntos de dados sejam valiosos para pesquisa, monitoramento de saúde pública e formulação e acompanhamento de políticas governamentais, eles podem apresentar riscos significativos à privacidade dos cidadãos. De fato, como os dados são disponibilizados publicamente, qualquer agente com acesso legítimo aos mesmos pode, em princípio, explorar vulnerabilidades na proteção de privacidade dos dados e, consequentemente, causar danos aos cidadãos. Um exemplo simples de tal risco seria um empregador inescrupuloso que discriminasse candidatos a uma vaga de emprego baseando-se em informações sensíveis sobre a saúde dos mesmos que pudessem ser inferidas a partir dos dados do sistema DATASUS. Como mencionado anteriormente, este risco de inferência de atributos sensíveis através de acesso legítimo aos dados é bem documentado na literatura e sua viabilidade já foi verificada em diversas bases, incluindo as dos Censos Educacionais do Brasil [Alvim et al. 2022].

Uma base de particular relevância na plataforma DATASUS é a SIA-PA (*Sistema de Informações Ambulatoriais - Produção Ambulatorial*). Esta base contém microdados (i.e., dados em nível individual) de dezenas de milhões de cidadãos que passaram por procedimentos ambulatoriais em estabelecimentos associados ao SUS desde 1994 até os dias atuais.[5] Neste trabalho avaliamos o risco de *inferência de atributos sensíveis* no SIA-PA do DATASUS. Mais precisamente, quantificamos o quanto um usuário com acesso legítimo ao sistema DATASUS pode inferir atributos sensíveis (como condição médica, enfermidades, ou procedimentos realizados) de titulares dos microdados na base. Assumimos que o usuário tem conhecimento a respeito de QIDs dos indivíduos escolhidos como alvo, isto é, atributos que não identificam o indivíduo por si próprios, mas que, quando combinados, permitem uma reidentificação com alta acurácia (como idade, sexo e município de residência) [Sweeney 2000]. Usando estes QIDs, o usuário pode filtrar a base de dados procurando o indivíduo alvo e, com isso, inferir seus atributos sensíveis com maior ou menor grau de certeza.

Como já mencionado, os ataques à privacidade podem ser classificados em ataques de pertencimento, de reidentificação, e de inferência de atributo sensível, sendo este último o foco do presente trabalho. Ataques de pertencimento não são particularmente danosos no contexto considerado, uma vez que a simples inferência de que um indivíduo está incluído na base de procedimentos ambulatoriais não causa danos imedi-

[2] `https://www.gov.br/inep/pt-br/assuntos/noticias/institucional/nota-de-esclarecimento-divulgacao-dos-microdados`

[3] Mais precisamente, em seu Artigo 12º, a LGPD estabelece que dados anonimizados (i.e., dados em que identificadores óbvios de indivíduos como nome completo e CPF, não estão presentes) não são considerados dados pessoais, exceto quando a anonimização puder ser revertida com esforços razoáveis. Para tanto, deve-se considerar fatores objetivos como o custo e o tempo necessários para reverter o processo de anonimização dadas as tecnologias disponíveis e desconsiderando-se o uso de meios de terceiros.

[4] `https://datasus.saude.gov.br/`

[5] `https://ces.ibge.gov.br/base-de-dados/metadados/ministerio-da-saude/sistema-de-informacoes-ambulatoriais-do-sus-sia-sus`

atos. A reidentificação do indivíduo, entretanto, pode causar danos, uma vez que se um registro é ligado a um titular de dados, inferem-se imediatamente os atributos sensíveis do titular presentes naquele registro. Destacamos, entretanto, que não é necessário haver reidentificação explícita de um indivíduo para ocorrer inferência de seus atributos sensíveis. Isto ocorre, por exemplo, quando o adversário consegue usar os QIDs de um alvo para filtrar a base a um número reduzido de registros compatíveis, em que todos contêm, digamos, um procedimento ambulatorial considerado de maior gravidade: neste caso, mesmo que o adversário não consiga saber exatamente qual o procedimento realizado pelo alvo, é possível ter certeza de que trata-se de um procedimento grave. Nos experimentos apresentamos tanto uma análise minuciosa dos riscos à privacidade para as bases de dados de todos os estados brasileiros do ano de 2023, quanto uma análise temporal do risco para estados selecionados do ano de 2019 até 2023.

As principais contribuições deste trabalho são portanto:

1. Um modelo formal para riscos de inferência de atributos sensíveis, em que titulares de dados na base podem ter valores de atributos considerados privados (e.g., dados médicos) inferidos por alguém com acesso legítimo e autorizado à base. Nossa metodologia é baseada em um adversário Bayesiano que utiliza a base de dados para atualizar seu conhecimento sobre o atributo sensível do alvo.
2. Uma bateria de experimentos que demonstram sérios riscos à privacidade nos microdados de procedimentos ambulatoriais do DATASUS, que contém dados individuais de dezenas de milhões de cidadãos brasileiros ao longo das últimas décadas, e que é protegida apenas pela técnica de desidentificação. Mais precisamente, quantificamos o impacto que diferentes níveis de conhecimento do adversário (i.e., quantidade e tipos de QIDs) causam no sucesso esperado de ataques de inferência de atributos. Os resultados também demonstram que, entre os diferentes estados brasileiros, há uma variabilidade considerável no sucesso esperado dos ataques.

Acreditamos que este trabalho é uma contribuição relevante para a análise de se a plataforma DATASUS encontra-se em conformidade com a LGPD. Além disso, este trabalho apresenta indicações de quais são os pontos mais vulneráveis em termos de privacidade (i.e., quais QIDs levam aos maiores riscos de inferência de atributo sensíveis), o que pode guiar o desenho de um melhor compromisso entre privacidade e utilidade na base de dados analisada.

Organização deste documento. O restante deste artigo está organizado da seguinte forma. Na Seção 2 referenciamos os principais trabalhos relacionados ao controle de divulgação estatística e à análise de riscos à privacidade. Na Seção 3 apresentamos a formalização do ataque de inferência de atributos em uma base de microdados. Na Seção 4 apresentamos a análise quantitativa experimental dos riscos de inferência de atributo sensível na base de procedimentos ambulatoriais do DATASUS. Por fim, na Seção 5 apresentamos nossas considerações finais e discutimos possíveis direções de trabalhos futuros.

2. Trabalhos Relacionados

Na literatura sobre controle de divulgação estatística é amplamente conhecido o compromisso (*trade-off*) entre privacidade e utilidade [Dinur and Nissim 2003, Dwork 2011, Fung et al. 2010, Hundepool et al. 2012]: uma maior garantia de privacidade aos titulares dos dados é geralmente acompanhada de uma menor utilidade dos dados para os analistas, e vice-versa. De acordo com trabalhos recentes, como em [Sarmin et al. 2024], não existe um método de mitigação de riscos à privacidade que se sobreponha aos demais a respeito do nível de utilidade provido, em particular, os autores mostram que este *trade-off* em geradores de dados sintéticos atuais não é melhor que aquele provido por métodos de anonimização tradicionais. Entre os desafios de pesquisa centrais desta área, temos a caracterização deste compromisso (e.g., estudos que quantificam a perda de privacidade por unidade de utilidade) e o desenvolvimento de métodos de controle de divulgação estatística que garantam níveis aceitáveis de privacidade e utilidade. Este compromisso é quantificado em [Abowd et al. 2021], onde os autores apresentam resultados teóricos demonstrando que, para garantir um nível mínimo de privacidade, á acurácia é degradada por um fator logarítmico. Este trabalho se concentra em quantificar o quão vulnerável é um conjunto de dados desidentificado em termos de ataque de inferência de atributos.

Nossa formalização de ataques segue o modelo proposto por Nunes et al. [Nunes 2021, Alvim et al. 2022], mas difere do mesmo nas premissas sobre o conhecimento a priori do adversário. Enquanto Nunes et al. consideram um adversário (mais informado) que conhece os quaseidentificadores de todos os indivíduos presentes na base de dados (e.g., representando, por exemplo, o próprio governo), neste trabalho consideramos um adversário menos informado (porém mais realista), que conhece os quaseidentificadores apenas de um alvo específico. Assim como em Nunes et al., quantificamos o sucesso esperado do adversário em inferir corretamente atributos sensíveis dos alvos.

Dentre os métodos que tentam explicar riscos à privacidade em termos práticos, o arcabouço matemático formal de fluxo de informação quantitativo (*quantitative information flow*) QIF [Alvim et al. 2020a] oferece um conjunto de princípios sólidos baseados em teoria da informação e teoria da decisão para raciocinar sobre o fluxo de informação sensível através de um sistema. O arcabouço de QIF apresenta um vasto conjunto de ferramentas, definições e propriedades que permitem modelar cenários de divulgação estatística explicáveis em linguagem relativamente acessível para a sociedade em geral. Alguns trabalhos recentes [Nunes 2021, Jurado et al. 2023, Athanasiou et al. 2024, Fernandes et al. 2024, Alvim et al. 2020b] utilizam QIF para modelar cenários de divulgação estatística com o objetivo de quantificar e explicar, em termos simples, as vulnerabilidades envolvidas nessas publicações de dados.

3. Formalização do ataque de inferência de atributo sensível

Nesta seção definimos formalmente o risco de inferência de atributo sensível sobre uma publicação de uma base de microdados (i.e., dados em nível de indivíduos) protegida apenas pelo método de desidentificação, em que se removem possíveis identificadores individuais óbvios dos registros (como nome, CPF, RG).

Modelo de adversário. O *adversário* é uma pessoa ou entidade que possui acesso legítimo e autorizado à base de dados e pode vir a inferir informações sensíveis so-

bre indivíduos representados na mesma. Existem, naturalmente, adversários mal-intencionados: por exemplo, um responsável pela contratação de novos funcionários em uma empresa pode usar uma base de dados do DATASUS para inferir o estado de saúde de candidatos e rejeitá-los com base nas informações obtidas. Entretanto, o adversário não é necessariamente malicioso: um pesquisador ou demógrafo idôneo que utilize a base de dados para fins legítimos pode, mesmo que inadvertidamente, inferir informações sensíveis sobre indivíduos, configurando, assim, uma violação de privacidade.

Assumimos que o adversário possui um único alvo, selecionado aleatoriamente na base com probabilidade uniforme, e tem o objetivo de inferir o valor de um atributo sensível deste alvo. O adversário sabe que o alvo está representado em um registro da base de dados e que este registro contém o valor do atributo considerado sensível. Assumimos também que o adversário pode obter os valores dos quaseidentificadores (QIDs) para o alvo escolhido e usar estes QIDs para filtrar os registros na base. Naturalmente, conforme aumenta o número de QIDs do alvo que o adversário conhece, o número de indivíduos com aqueles valores específicos tende a se tornar cada vez menor, até o limite em que o indivíduo se torna único. Avaliamos então a probabilidade esperada do adversário identificar corretamente o valor sensível do alvo, dado seu conhecimento sobre variadas combinações de possíveis quaseidentificadores do mesmo.

Para formalizar o adversário descrito acima, começamos por introduzir uma notação específica.

- Seja $A = \{a_1, \ldots, a_m\}$ um conjunto (finito) de *atributos* (ou *colunas*) de interesse em um determinado cenário. Para cada $1 \leq i \leq m$, denotamos por $\texttt{dom}(a_i)$ o domínio de a_i.
- Um *registro* (ou *linha*) x é um mapeamento de atributos para valores, isto é, $x = \langle x[a_1], \ldots, x[a_m] \rangle$, em que $x[a_i] \in \texttt{dom}(a_i)$ é o valor assumido pelo atributo a_i no registro x. O domínio de todos os registros possíveis sobre o conjunto de atributos A é denotado por $\texttt{dom}(A) = \texttt{dom}(a_1) \times \ldots \times \texttt{dom}(a_m)$.
- Uma *base de dados* sobre um conjunto de atributos A é um multiconjunto (finito) $D = \{\!\{x_1, x_2, \ldots, x_n\}\!\}$ de registros, cada um pertencente a $\texttt{dom}(A)$.
- Dado um subconjunto de atributos $A' \subseteq A$, denotamos por $x[A']$ o *sub-registro* (ou *sub-linha*) de x, que consiste na projeção de x sobre A', ou seja, a subtupla de x contendo apenas os valores correspondentes aos atributos em A'. Os sub-registros correspondentes aos atributos em $A' = \{a_{i_1}, \ldots, a_{i_k}\}$ têm como domínio o conjunto $\texttt{dom}(A') = \texttt{dom}(a_{i_1}) \times \ldots \times \texttt{dom}(a_{i_k})$.

O *conhecimento do adversário* é formalizado como a seguir.

- O adversário conhece o número de registros $|D|$ presentes na base de dados D de interesse e também a distribuição do atributo sensível $a_s \in A$ de interesse em D.
- Dado um registro $x \in D$ escolhido como alvo, o adversário consegue obter o valor dos QIDs $x[Q]$ deste alvo, para um conjunto de QIDs $Q \subseteq A$ de interesse.

Modelo de ataque de inferência de atributo sensível. Em um *ataque de inferência de atributo sensível*, o adversário busca inferir o valor $x[a_s]$ do atributo sensível $a_s \in A$ pertencente a um indivíduo alvo x. O sucesso do ataque é formalizado como a seguir.

- A *vulnerabilidade a priori* representa a probabilidade esperada de um adversário ótimo inferir corretamente o valor $x[a_s]$ do atributo sensível do alvo x na base D sem acesso aos QIDs $x[Q]$ do mesmo. A estratégia ótima neste caso é considerar a distribuição do atributo sensível na base D como um todo e selecionar o valor mais frequente – se houver dois ou mais valores sensíveis com a maior frequência, o adversário escolhe aleatoriamente um entre eles. A vulnerabilidade a priori $prior(D)$ de um ataque de inferência de atributo é formalizada como

$$prior(D) \stackrel{\text{def}}{=} \max_{v \in \mathtt{dom}(a_s)} \frac{|x \in D : x[a_s] = v|}{|D|}, \tag{1}$$

  onde $\max\limits_{v \in \mathtt{dom}(a_s)} |x \in D : x[a_s] = v|$ é o número de registros com o valor mais comum para o atributo sensível a_s na base D, e $|D|$ é o número total de registros em D.
- A *vulnerabilidade a posteriori* representa a probabilidade esperada de um adversário ótimo inferir corretamente o valor $x[a_s]$ do atributo sensível do alvo x na base D tendo conhecimento dos valores $x[Q]$ dos QIDs do mesmo. O ataque ocorre como a seguir.
  Primeiramente o adversário filtra na base D os registros que compartilham os mesmos valores $x[Q]$ de QIDs que o alvo. A estratégia ótima para este adversário é selecionar o valor sensível mais frequente entre os registros filtrados como sua estimativa para o valor sensível $x[a_s]$ do alvo – se houver dois ou mais valores sensíveis com a maior frequência, o adversário escolhe aleatoriamente um entre eles. Dados um registro alvo x fixo, a vulnerabilidade a posteriori $success(x)$ do registro x é formalizada como

$$success(x) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{|\text{smf}(x)|} & \text{, se } x[a_s] \in \text{smf}(x) \\ 0 & \text{, caso contrário,} \end{cases} \tag{2}$$

  onde $\text{smf}(x)$ é o conjunto dos atributos sensíveis mais frequentes entre os candidatos que compartilham os mesmos QIDs $x[Q]$ do alvo.
  A vulnerabilidade a posteriori geral é definida como a expectativa da probabilidade de sucesso assumida sobre a distribuição de probabilidade de que cada registro no banco de dados original (D) tenha sido selecionado como alvo -– que assumimos ser uniforme, i.e., $\mathsf{Pr}[x] = 1/|D|$. Sendo assim, a vulnerabilidade a posteriori de um ataque de inferência de atributo $posterior(D)$ é definida como

$$\begin{aligned} posterior(D) &\stackrel{\text{def}}{=} \sum_{x \in D} \mathsf{Pr}[x] success(x). \\ &= \frac{1}{|D|} \sum_{x \in D} success(x). \end{aligned} \tag{3}$$

- A *degradação da privacidade* (i.e., o *vazamento de informação)* do ataque é uma comparação entre as vulnerabilidades a priori e a posteriori, indicando o quanto o acesso aos QIDs aumenta, em expectativa, a probabilidade de sucesso do adversário no ataque, caso o alvo seja escolhido uniformemente na base. Em particular, o *vazamento aditivo* de um ataque de inferência de atributo de uma base de dados D é formalizado como

$$\mathcal{L}^{+}(D) \stackrel{\text{def}}{=} posterior(D) - prior(D). \tag{4}$$

4. Avaliação quantitativa experimental

Nesta seção descrevemos uma bateria de experimentos que avaliam os potenciais riscos à privacidade nos microdados de procedimentos ambulatoriais do DATASUS. O código fonte utilizado nos experimentos contendo a implementação dos ataques descritos na Seção 3 está disponível publicamente no GitHub [6].

4.1. Bases de dados utilizadas

A plataforma DATASUS possui várias bases de dados disponíveis publicamente, com variadas características e estruturas de organização. Neste trabalho focamos na base do *Sistema de Informação Ambulatorial - Produção Ambulatorial* (SIA-PA),[7] que contém microdados de atendimentos ambulatoriais realizados no país nas últimas décadas. A base SIA-PA contém 60 atributos por registro,[8] dos quais identificamos oito como potenciais quaseidentificadores (i.e., atributos cujos valores um adversário razoavelmente bem-informado pode facilmente obter para um indivíduo escolhido como alvo), conforme a Tabela 1(a), e um como potencial atributo sensível, conforme a Tabela 1(b). Os demais 51 atributos não foram selecionados como candidatos a QIDs ou a atributos sensíveis neste estudo pelos seguintes motivos: (i) não estarem relacionados ao titular dos dados, (ii) serem redundantes com respeito a outros atributos já considerados, ou (iii) serem atributos que possuem valores inválidos para um número alto de registros.

Nos experimentos desta seção, usamos as bases de dados do SIA-PA de algumas UFs do Brasil entre 2019 e 2023 (na análise comparativa ao longo dos anos) e as bases de dados de todas as UFs para o ano de 2023 (nas análises mais minuciosas restritas a um único ano). Em todas as bases utilizadas fizemos um pré-processamento para remover todos os registros que continham valores inválidos para qualquer atributo. A maioria dos registros foi preservada nas bases de dados de 2023 após o pré-processamento, com a base do Acre tendo a menor proporção de registros preservados (74.77%), e a base do Amazonas tendo a maior proporção de registros preservados (91.47%). Nota-se uma diminuição considerável do número de registros que contêm valores inválidos para qualquer atributo de 2019 para 2023, e.g., 57,78% dos registros do Distrito Federal em 2019 eram inválidos, enquanto que em 2023 somente 11,14% dos registros eram inválidos. Ressaltamos que, da totalidade de registros inválidos em todas as bases de dados desta análise comparativa, a variável `PA_RACACOR` possuía um valor inválido em 99,9% dos registros.

4.2. Cenários de ataque

Conforme descrito na Seção 3, assumimos que dado um indivíduo alvo x, o adversário conhece os valores $x[Q]$ de seus atributos quaseidentificadores $Q \subseteq A$. Como mencionado anteriormente, à medida em que aumenta o número de QIDs do alvo que o adversário

[6] `https://github.com/ramongonze/privattacks`

[7] As bases de dados utilizadas neste trabalho foram obtidas de um servidor FTP `ftp://ftp.datasus.gov.br/dissemin/publicos/SIASUS/200801_/Dados/`. Para facilitar o acesso aos dados, foi usado o seguinte pacote R para baixar os conjuntos de dados: `https://github.com/rfsaldanha/microdatasus`.

[8] A descrição dos atributos está disponível no relatório técnico [Ministério da Saúde 2019] da *Divisão de Análise e Administração de Dados* (DIAAD).

[9] A tabela contendo o código e a descrição de todos os procedimentos ambulatoriais está disponível em `http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp`.

Atributo	Descrição
PA_SEXO	Sexo do paciente.
PA_RACACOR	Raça/Cor do paciente: 01 - Branco, 02 - Preta, 03 - Parda, 04 - Amarela, 05 - Indígena, 99 - Sem informação.
PA_IDADE	Idade do paciente em anos.
PA_MUNPCN	Código do município de residência do paciente.
PA_CODUNI	Código do Estabelecimento no CNES (Cadastro Nacional de Estabelecimentos de Saúde).
PA_MVM	Data de processamento (AAAAMM).
PA_CMP	Data de realização do procedimento (AAAAMM).
UF (implícito)	Unidade da federação onde ocorreu o procedimento.

(a) Atributos selecionados como candidatos a quaseidentificadores. Não há um atributo específico para a unidade da Federação (UF) onde ocorreu o procedimento ambulatorial; entretanto, como as bases de dados disponíveis para *download* são divididas por UFs, podemos considerar o atributo como implicitamente disponível.

Atributo	Descrição
PA_PROC_ID	Código do Procedimento Ambulatorial [6]

(b) Atributo selecionado como candidato a atributo sensível.

Tabela 1. Lista de atributos selecionados nas bases de dados SIA-PA.

conhece, o número de indivíduos com aqueles valores específicos tende a se tornar cada vez menor, até o limite em que o indivíduo se torna único. Neste sentido, avaliamos o risco à privacidade para adversários que tenham como conhecimento auxiliar cada uma das possíveis combinações (não-vazias) de candidatos a quaseidentificadores descritos na Tabela 1(a), perfazendo um total de $2^7 - 1 = 127$ cenários distintos (considerando que UF esteja sempre disponível).

Exemplo real de ataque exitoso. Descrevemos agora um ataque real de inferência de atributo sensível na base do estado de Minas Gerais (MG) do ano de 2023. Para efeito de privacidade, alteramos ou omitimos alguns atributos que permitam sua reprodução. Consideramos dois indivíduos, Ana e Bruno (nomes fictícios), residentes no município M (nome fictício). Bruno, o empregador de Ana, sabe que em fevereiro de 2023 ela foi a Belo Horizonte realizar um procedimento no Hospital H (nome também fictício) mas não sabe qual procedimento. Para descobrir o motivo da visita de Ana ao hospital, Bruno filtra a base de dados publicamente disponível da SIA-PA para Minas Gerais em 2023 usando os seguintes atributos quaseidentificadores que ele conhece sobre Ana:

- PA_MUNPCN: 000000 (Município M)
- PA_SEXO: F (Feminino)
- PA_IDADE: 45
- PA_RACACOR: 1 (Branca)
- PA_CMP: 202302 (Fevereiro 2023)
- PA_CODUNI: 0000000 (identificador único do Hospital H)

Filtrando a base por esta combinação de QIDs, Bruno encontra um único registro e conclui que o mesmo deve pertencer a Ana. Examinando o atributo PA_PROC_ID, ele

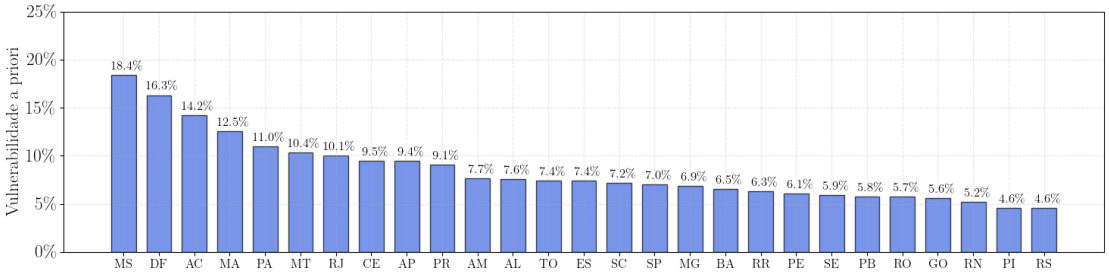


Figura 1. Vulnerabilidade a priori para todas as bases do SIA-PA de 2023 em relação ao atributo sensível `PA_PROC_ID` (Código do Procedimento Ambulatorial). Os valores representam a probabilidade de o adversário inferir corretamente o valor do atributo de um indivíduo aleatoriamente selecionado como alvo, sem ter acesso aos QIDs do mesmo.

obtém o código do procedimento ambulatorial `0304020346`, que corresponde a uma "*Hormonioterapia do Carcinoma de Mama Avançado – 1ª Linha*", que consiste em um procedimento inicial realizado em pacientes diagnosticadas com câncer de mama. A partir dessas informações, Bruno infere com sucesso que Ana tem câncer de mama.

O tipo de atributo sensível do exemplo acima não é inusual: existem diversos outros procedimentos sensíveis que podem ser descobertos com a base SIA-PA, como se uma pessoa realizou um aborto (código 0409060070 - ESVAZIAMENTO DE UTERO POS-ABORTO POR ASPIRACAO MANUAL INTRA-UTERINA (AMIU)) ou se ela possui HIV/AIDS (código 0303180072 - TRATAMENTO DE HIV/AIDS). O vazamento de informações desta natureza pode ser prejudicial para os detentores de dados e ocasionar discriminações em diferentes cenários.

4.3. Quantificação do risco esperado de ataques de inferência de atributo sensível

O exemplo da seção anterior ilustra um ataque de inferência de atributo sensível exitoso em uma base de dados real do DATASUS. Em princípio, pode-se ter a impressão de que o sucesso em um ataque deste tipo seria uma ocorrência rara, e que Ana deve ter sido particularmente desafortunada dentro da base de dados. Aqui, entretanto, verificamos que o caso de Ana não é incomum no sistema DATASUS e quantificamos a expectativa de indivíduos aleatoriamente selecionados como alvos terem seus atributos sensíveis inferidos. A nossa análise considera separadamente cada UF e é focada nas bases de dados de 2023.

Análise da vulnerabilidade a priori. Nossa análise começa pela computação da vulnerabilidade a priori de cada base em relação ao atributo sensível `PA_PROC_ID` (Código do Procedimento Ambulatorial). Definida na Eq. 1, esta vulnerabilidade representa o sucesso esperado do adversário ao tentar inferir o valor sensível do alvo $x[a_s]$ sem usar nenhuma informação auxiliar de QIDs. A Figura 1 apresenta a vulnerabilidade a priori computada em cada UF. Podemos interpretar, por exemplo, $prior(\texttt{MS})$ como "*A chance esperada do adversário adivinhar corretamente o procedimento ambulatorial de um alvo no estado do Mato Grosso do Sul (MS), sem usar nenhuma informação auxiliar (QIDs) sobre o alvo, é de aproximadamente 18,4%.*"

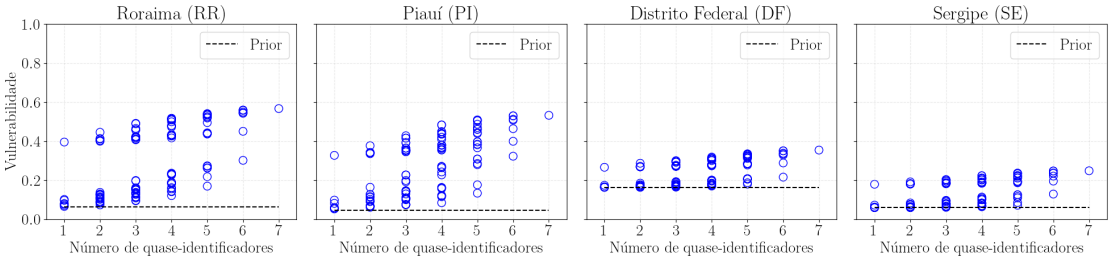


Figura 2. Vulnerabilidade a priori e a posteriori para algumas UFs selecionadas nas bases do SIA-PA de 2023 em relação ao atributo sensível `PA_PROC_ID` (Código do Procedimento Ambulatorial). Cada ponto representa a vulnerabilidade a posteriori de uma combinação de QIDs usada como informação auxiliar pelo adversário, e a posição do ponto no eixo vertical indica a vulnerabilidade a posteriori do ataque correspondente.

Análise da vulnerabilidade a posteriori e do vazamento aditivo. Em seguida prosseguimos com a computação da vulnerabilidade a posteriori de cada base em relação ao mesmo atributo sensível. Definida na Eq. 3, esta vulnerabilidade representa o sucesso esperado do adversário ao tentar inferir o valor sensível do alvo $x[a_s]$ já usando como informação auxiliar os QIDs do alvo. Subtraindo a vulnerabilidade a priori da posteriori, obtemos o vazamento aditivo (Eq. 4) que representa o quanto a execução do ataque aumenta a chance de sucesso do adversário em inferir o atributo sensível de um indivíduo aleatoriamente selecionado como alvo na base.

A Figura 2 apresenta as vulnerabilidades a priori e a posteriori para algumas UFs selecionadas. Considerando os candidatos a QIDs da Tabela 1(a), há $\binom{7}{n}$ combinações diferentes de n QIDs que o adversário poderia conhecer sobre o alvo em cada tabela (lembrando que o oitavo QID é a própria UF). Cada ponto na Figura 2 representa uma combinação possível de QIDs que o adversário pode ter como conhecimento auxiliar. A posição de cada ponto no eixo vertical representa a vulnerabilidade a posteriori para o ataque correspondente. Para o gráfico de Roraima (RR), por exemplo, vemos que combinações de 2 QIDs levam a uma vulnerabilidade a posteriori de quase 6% até cerca de 44%, enquanto a vulnerabilidade a priori é sempre de em torno de 6%, o que significa que a utilização de 2 QIDs pode levar a ataques com vazamento aditivo de cerca de 0% (= 6% a posteriori - 6% a priori) a até cerca de 38% (= 44% a posteriori - 6% a priori).

A Tabela 2 apresenta para cada UF os resultados das combinações de QIDs com maior vazamento aditivo. É possível observar que algumas UFs apresentam um vazamento alto para qualquer número de QIDs utilizados pelo adversário. Por exemplo, quando o adversário utiliza 7 QIDs, os três estados com maior vazamento são Roraima (50,33%), Piauí (48,6%) e Alagoas (43,92%). Por outro lado, os estados com os menores vazamentos são Sergipe (18,9%), Distrito Federal (19,08%) e Rio de Janeiro (22,95%). Note que uma alta vulnerabilidade a priori (i.e., esperança do sucesso do adversário antes de ter acesso à base de dados) não implica em alto vazamento (i.e., aumento na esperança do sucesso do adversário ao adquirir informação auxiliar). De fato, podemos observar que o Distrito Federal tem a segunda maior vulnerabilidade a priori (16,32%), mas tem o segundo menor vazamento para qualquer número de QIDs usados pelo adversário. O oposto também pode acontecer: Roraima tem uma baixa vulnerabilidade a priori (6,34%),

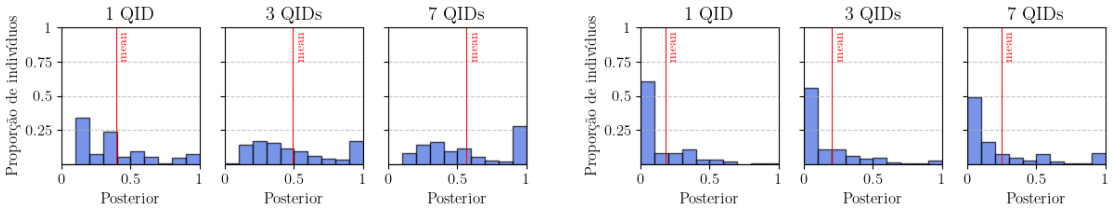
Estado	Prior	Vazamento Aditivo $\mathcal{L}^+$						
		1 QID	2 QIDs	3 QIDs	4 QIDs	5 QIDs	6 QIDs	7 QIDs
AC	14,24%	22,65%	28,59%	32,92%	37,16%	40,39%	43,04%	43,33%
AL	7,61%	28,17%	33,71%	36,53%	39,80%	41,77%	43,66%	43,92%
AP	9,45%	19,30%	23,31%	27,47%	30,36%	32,69%	34,70%	35,38%
AM	7,66%	17,67%	19,54%	22,12%	23,16%	24,06%	24,77%	25,02%
BA	6,53%	25,83%	28,78%	31,31%	34,51%	36,60%	38,33%	38,77%
CE	9,51%	18,60%	21,09%	22,92%	25,35%	26,73%	27,85%	27,99%
DF	16,32%	10,26%	12,34%	13,65%	15,44%	17,12%	18,71%	19,08%
ES	7,40%	23,58%	27,54%	30,17%	33,71%	36,39%	38,90%	39,19%
GO	5,63%	23,09%	26,36%	29,55%	33,74%	36,09%	38,19%	38,46%
MA	12,53%	20,95%	24,71%	28,28%	31,24%	33,41%	35,29%	35,65%
MG	6,86%	21,03%	22,92%	24,95%	27,14%	28,83%	30,38%	30,76%
MT	10,38%	19,19%	21,21%	24,36%	26,03%	27,41%	28,73%	28,84%
MS	18,43%	22,98%	26,13%	29,02%	31,17%	33,52%	35,39%	35,56%
PA	10,97%	20,81%	23,34%	25,96%	28,34%	29,96%	31,17%	31,34%
PB	5,76%	24,25%	27,29%	30,60%	34,87%	36,98%	38,83%	39,24%
PR	9,10%	26,24%	29,53%	31,95%	34,92%	36,96%	38,98%	39,14%
PE	6,06%	23,42%	26,78%	30,14%	34,22%	36,37%	38,28%	39,33%
PI	4,61%	28,08%	33,04%	38,12%	43,67%	46,22%	48,35%	48,60%
RJ	10,06%	14,30%	16,64%	18,49%	20,43%	21,70%	22,62%	22,95%
RN	5,21%	22,97%	26,03%	29,30%	33,43%	35,53%	37,27%	37,43%
RS	4,55%	19,87%	21,34%	23,06%	25,32%	26,67%	27,52%	27,88%
RO	5,75%	24,68%	26,93%	29,79%	31,77%	33,77%	35,29%	35,66%
RR	6,34%	33,18%	38,18%	42,76%	45,38%	47,76%	49,61%	50,33%
SC	7,17%	30,03%	32,67%	35,96%	38,23%	40,25%	41,51%	41,76%
SP	7,00%	25,05%	28,78%	31,44%	34,15%	36,88%	39,51%	40,04%
SE	5,89%	11,98%	13,22%	14,39%	16,36%	17,68%	18,79%	18,90%
TO	7,44%	28,51%	32,43%	36,41%	39,08%	40,86%	42,52%	42,59%

Tabela 2. Vulnerabilidade a priori e vazamento aditivo de todas as UFs das bases do SIA-PA de 2023 em relação ao atributo sensível `PA_PROC_ID` (Código do Procedimento Ambulatorial). Em cada coluna, os 3 valores mais altos e os 3 mais baixos são destacados em laranja e azul, respectivamente.

mas o maior vazamento para qualquer número de QIDs.

Para uma análise mais profunda, a Figura 3 apresenta a distribuição da vulnerabilidade a posteriori para os estados de Roraima e Sergipe, os de maior e menor vazamento, respectivamente. Podemos observar que a variabilidade da vulnerabilidade é alta, o que significa que alguns registros nas bases de dados são muito comuns em relação aos valores dos QIDs (levando à uma vulnerabilidade a posteriori baixa) e alguns registros são bastante únicos (o que leva à uma vulnerabilidade a posteriori alta). Para a base de dados de Roraima, utilizando 7 QIDs, o adversário é capaz de adivinhar corretamente, com uma precisão de mais de 90%, o procedimento ambulatorial de mais de 28% dos registros (que corresponde a 83.154 registros). Os histogramas mostram a distribuição para a combinação de 1, 3 e 7 QIDs que produziram a vulnerabilidade a posteriori máxima.

Para explicar a disparidade no vazamento entre estados, investigamos correlações entre o vazamento de informação e variáveis possivelmente explanatórias. Um equívoco comum em discussões sobre privacidade é assumir a hipótese de "*privacidade na multidão*", ou seja, esperar que indivíduos em grandes populações estejam inerentemente protegidos. Os resultados apresentados na Tabela 2 demonstram que esta hipótese é re-



(a) Roraima (RR). **(b) Sergipe (SE).**

Figura 3. Distribuição da vulnerabilidade a posteriori usando 1, 3 e 7 quasi-identificadores. Foi selecionada a combinação de QIDs que produz a vulnerabilidade a posteriori máxima.

almente inválida nas bases analisadas: ao usar 7 QIDs, o vazamento para as bases do Acre (879.370 registros) e São Paulo (112.637.066 registros) são 43,33% e 40,04%, respectivamente. Por outro lado, o vazamento para as bases do Distrito Federal (5.199.384 registros) e Rio de Janeiro (71.712.152 registros) são significativamente menores, 19,08% e 22,95%, respectivamente. O Coeficiente de Correlação de Pearson do vazamento aditivo vs. número de registros, considerando o número de QIDs de 1 a 7, está no intervalo de -0,19 a -0,09, o que sugere uma correlação muito fraca.

Uma segunda análise interessante diz respeito à entropia do atributo sensível. A entropia mede a incerteza ou diversidade dos valores do atributo sensível, refletindo o quão difícil pode ser para um invasor inferir corretamente o valor do atributo. Ao examinar se atributos com entropia mais alta realmente levam a vazamentos menores, podemos entender como a distribuição de dados afeta diretamente a vulnerabilidade da privacidade. Semelhante ao tamanho da base de dados, identificamos que não há correlação direta forte entre o vazamento aditivo e a entropia, com o Coeficiente de Correlação de Pearson variando de 0,21 a 0,28. Estes números corroboram resultados conhecidos na literatura sobre QIF de que entropia nem sempre representa uma medida acurada de vulnerabilidade a ataques e, portanto, deve ser usada com cautela como medida de privacidade [Smith 2009].

São necessários trabalhos futuros para compreender a disparidade de vazamento entre os estados. Uma possível direção a ser investigada é se há uma diferença na correlação dos QIDs entre os estados, que é calculada a partir da distribuição conjunta destes. Considerando que o ataque do adversário inclui filtrar os registros que possuem os mesmos valores dos QIDs do alvo, a distribuição conjunta destes ditará o tamanho do subconjunto de registros obtidos pelo adversário como possíveis candidatos ao alvo.

Impacto dos QIDs utilizados. É sabido que certos atributos QIDs particionam registros em grupos menores de forma mais eficaz do que outros, influenciando diretamente a eficácia dos ataques de inferência de atributos. Analisamos como a remoção de cada um dos sete QIDs listados na Tabela 1(a) afeta o vazamento. Especificamente, a Tabela 3 apresenta o vazamento máximo atingível por um adversário quando um desses QIDs não está disponível na base de dados. O atributo com maior impacto no vazamento em todos os conjuntos de dados é claramente `PA_CODUNI`, seguido pela idade do indivíduo (`PA_IDADE`). Naturalmente, filtrar o conjunto de registros restringindo-o a um único estabelecimento onde foi realizado o procedimento ambulatorial, reduz consideravelmente o

número de registros que podem pertencer ao alvo. O mesmo raciocínio pode ser aplicado à idade. Isto explica o alto impacto destes atributos no vazamento de informação.

Estado	Vazamento máx. c/ todos os QIDs	Vazamento máximo após remover um QID						
		PA_CODUNI	PA_IDADE	PA_MUNPCN	PA_RACACOR	PA_SEXO	PA_MVM	PA_CMP
AC	57,57%	**30,01%**	**42,83%**	**53,12%**	54,33%	54,91%	57,28%	57,28%
AL	51,53%	**28,86%**	**40,75%**	**47,83%**	49,66%	49,57%	51,10%	51,26%
AP	44,82%	**28,20%**	**33,95%**	42,70%	42,89%	**41,99%**	44,15%	43,98%
AM	32,67%	**15,79%**	**28,70%**	31,73%	31,95%	**31,68%**	32,43%	32,43%
BA	45,30%	**25,98%**	**36,89%**	**41,94%**	43,19%	43,58%	44,78%	44,86%
CE	37,50%	**24,53%**	**30,81%**	**34,10%**	36,40%	36,09%	37,35%	37,36%
DF	35,40%	**21,53%**	**28,80%**	**32,87%**	33,41%	33,80%	35,01%	35,03%
ES	46,58%	**26,38%**	**35,34%**	**42,00%**	43,75%	44,10%	46,29%	46,27%
GO	44,09%	**23,76%**	**33,74%**	**38,76%**	41,77%	42,01%	43,81%	43,75%
MA	48,17%	**29,28%**	**38,13%**	**45,16%**	46,33%	45,97%	47,82%	47,81%
MG	37,62%	**22,00%**	**30,75%**	**34,84%**	35,83%	36,07%	37,23%	37,22%
MT	39,21%	**25,24%**	**32,82%**	37,84%	**37,36%**	37,90%	39,11%	39,10%
MS	53,99%	**34,47%**	**44,59%**	**51,48%**	51,59%	52,12%	53,82%	53,82%
PA	42,32%	**24,44%**	**35,17%**	**40,00%**	41,12%	40,70%	42,15%	42,08%
PB	45,00%	**27,06%**	**34,92%**	**39,90%**	43,00%	43,15%	44,59%	44,57%
PR	48,24%	**26,61%**	**38,34%**	**44,07%**	46,22%	46,11%	48,08%	48,04%
PE	45,39%	**24,32%**	**34,93%**	**40,84%**	43,55%	43,30%	44,16%	44,34%
PI	53,22%	**32,25%**	**39,91%**	**46,37%**	51,20%	50,76%	52,96%	52,93%
RJ	33,01%	**17,31%**	**27,04%**	32,10%	**30,66%**	31,67%	32,68%	32,62%
RN	42,64%	**25,73%**	**32,88%**	**38,21%**	40,59%	40,92%	42,47%	42,47%
RS	32,43%	**17,70%**	**26,72%**	**29,71%**	31,57%	31,02%	32,07%	32,07%
RO	41,41%	**27,49%**	**33,95%**	39,50%	**39,24%**	39,91%	41,02%	41,04%
RR	56,67%	**30,14%**	**45,01%**	54,38%	54,86%	**54,27%**	55,95%	55,71%
SC	48,93%	**26,82%**	**40,23%**	**46,18%**	47,68%	46,90%	48,68%	48,61%
SP	47,04%	**22,51%**	**35,63%**	**43,24%**	44,01%	44,42%	46,51%	46,52%
SE	24,79%	**12,85%**	**19,50%**	**22,05%**	23,48%	23,69%	24,68%	24,67%
TO	50,02%	**33,54%**	**40,65%**	**47,66%**	48,30%	48,38%	49,95%	49,91%

Tabela 3. Vazamento máximo atingível por um adversário quando cada QID é removido do conjunto de dados. Valores mais altos indicam QIDs menos críticos.

Variabilidade da vulnerabilidade a posteriori ao longo dos anos. A Figura 4 mostra a evolução do vazamento ao longo de 5 anos, de 2019 a 2023, para os três estados com maior vazamento na Tabela 2 (Roraima - RR, Piauí - PI e Alagoas - AL) e os três com o menor vazamento (Rio de Janeiro - RJ, Sergipe - SE e Distrito Federal - DF). Houve uma diminuição sutil no vazamento para Roraima e Alagoas de 2019 a 2023, enquanto o Piauí teve um aumento sutil no vazamento. Por outro lado, todos os três estados com menos vazamento mantiveram um vazamento menor ao longo dos anos. Aparentemente, não houve mudança significativa no vazamento de 2019 a 2023.

5. Considerações finais

Neste trabalho, propomos um modelo formal para ataques de inferência de atributos que nos permite responder à pergunta: "*Qual é a probabilidade esperada de que um adversário, conhecendo alguns quaseidentificadores de um indivíduo aleatoriamente selecionado como alvo na base de dados, infira corretamente o valor de um atributo sensível?*".

Avaliamos o modelo usando os conjuntos de dados de procedimento ambulatorial do SIA-PA disponíveis online na plataforma DATASUS. Os resultados mostram um vazamento considerável. Para alguns estados brasileiros, há registros que estão sob o risco (vulnerabilidade a posteriori) de quase 60%, ou seja, um adversário na posse de alguns

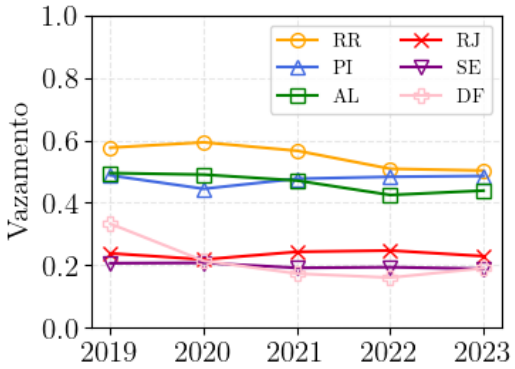


Figura 4. Análise da variabilidade do vazamento ao longo dos anos. Cada ponto representa o vazamento máximo alcançado utilizando todos os QIDs no ano em questão.

QIDs (e.g., idade, sexo, cidade de residência) pode inferir, com uma taxa de sucesso de quase 60%, o procedimento ambulatorial que seu alvo fez em uma determinada data.

A divulgação de dados estatísticos de saúde é essencial para propósitos como pesquisa e desenvolvimento e acompanhamento de políticas públicas. Entretanto, o estado de saúde dos cidadãos constitui informação privada que não deve ser publicamente acessível, inclusive de acordo com a legislação vigente. É reconhecido pela comunidade científica de que a percepção dos titulares de dados quanto à preservação da privacidade de suas informações afeta diretamente sua anuência e consentimento para uso dos dados, o que afeta, por conseguinte, a própria qualidade do conjunto de dados. Desta maneira, é imprescindível a proteção da privacidade dos titulares dos dados das bases do sistema DATASUS, sob pena, inclusive, de uma deterioração da visão do público sobre a mesma.

Neste sentido, aferir o nível de privacidade garantido por divulgações estatísticas é crucial para tranquilizar os titulares de dados e a sociedade como um todo (quando a privacidade é, de fato, preservada) ou para indicar ao gestor público vulnerabilidades em seu métodos de divulgação (quando há sinais de violações de privacidade). Um dos objetivos do presente estudo é alertar a comunidade sobre sérios riscos à privacidade identificados no sistema DATASUS. Esperamos que as vulnerabilidades experimentalmente verificadas possam ajudar a guiar os gestores públicos no desenvolvimento de soluções para mitigação de potenciais violações da LGPD.

Quanto a trabalhos futuros, identificamos duas principais direções. A primeira consiste em comparar o risco à privacidade existente em bases de dados de saúde brasileiras com o panorama internacional que adotam o mesmo formato que o SIA-PA (microdados desidentificados), como, por exemplo, a base de dados MIMIC [Johnson et al. 2023], disponibilizada pelo *Beth Israel Deaconess Medical Center* em Israel, e a base eICU [Pollard et al. 2018], divulgada pela *Philips Healthcare* nos Estados Unidos. Como segunda direção, consideramos testar diferentes técnicas de mitigação, como privacidade diferencial [Dwork et al. 2006], que podem ajudar a reduzir o risco de inferência de atributos, mantendo um nível razoável de utilidade.

Referências

Abowd, J., Ashmead, R., Cumings-Menon, R., Garfinkel, S., Kifer, D., Leclerc, P., Sexton, W., Simpson, A., Task, C., and Zhuravlev, P. (2021). An uncertainty principle is a price of privacy-preserving microdata. *Advances in neural information processing systems*, 34:11883–11895.

Alvim, M. S., Chatzikokolakis, K., McIver, A., Morgan, C., Palamidessi, C., and Smith, G. (2020a). *The Science of Quantitative Information Flow*. Information Security and Cryptography. Springer International Publishing, Cham, Switzerland.

Alvim, M. S., Fernandes, N., McIver, A., Morgan, C., and Nunes, G. H. (2022). Flexible and scalable privacy assessment for very large datasets, with an application to official governmental microdata. *Proc. Priv. Enhancing Technol.*, 2022(4):378–399.

Alvim, M. S., Fernandes, N., McIver, A., and Nunes, G. H. (2020b). On Privacy and Accuracy in Data Releases (Invited Paper). In Konnov, I. and Kovács, L., editors, *31st International Conference on Concurrency Theory (CONCUR 2020)*, volume 171 of *Leibniz International Proceedings in Informatics LIPIcs)*, pages 1:1–1:18, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Athanasiou, A., Chatzikokolakis, K., and Palamidessi, C. (2024). Self-defense: Optimal qif solutions and application to website fingerprinting. *arXiv preprint arXiv:2411.10059*.

Dinur, I. and Nissim, K. (2003). Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 202–210.

Dwork, C. (2011). A firm foundation for private data analysis. *Communications of the ACM*, 54(1):86–95.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer.

EU (2016). Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). Available at `https://eur-lex.europa.eu/eli/reg/2016/679/oj`.

Fernandes, N., McIver, A., and Sadeghi, P. (2024). Explaining epsilon in local differential privacy through the lens of quantitative information flow. In *2024 IEEE 37th Computer Security Foundations Symposium (CSF)*, pages 419–432. IEEE.

Fung, B. C., Wang, K., Fu, A. W.-C., and Philip, S. Y. (2010). *Introduction to privacy-preserving data publishing: Concepts and techniques*. Chapman and Hall/CRC.

Government of Australia (1988). Privacy Act 1988. `https://www.legislation.gov.au/Details/C2015C00598`.

Government of the United States of America (2002). Confidential information protection and statistical efficiency act (cipsea). `https://www.eia.gov/cipsea/cipsea.pdf`.

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Nordholt, E. S., Spicer, K., and De Wolf, P.-P. (2012). *Statistical disclosure control*, volume 2. Wiley New York.

Johnson, A. E., Bulgarelli, L., Shen, L., Gayles, A., Shammout, A., Horng, S., Pollard, T. J., Hao, S., Moody, B., Gow, B., et al. (2023). Mimic-iv, a freely accessible electronic health record dataset. *Scientific data*, 10(1):1.

Jurado, M., Alvim, M., Gonze, R., and Palamidessi, C. (2023). Analyzing the shuffle model through the lens of quantitative information flow. Technical report.

Matthews, G. J. and Harel, O. (2011). Data confidentiality: A review of methods for statistical disclosure limitation and methods for assessing privacy. *Statistics Surveys*, 5:1–29.

Ministério da Saúde (2019). Informe técnico - disseminação de dados em saúde - siasus. DIAAD - Divisão de Análise e Administração de Dados. Available at `ftp://ftp.datasus.gov.br/dissemin/publicos/SIASUS/200801_/Doc/Informe_Tecnico_SIASUS_2019_07.pdf`.

Nunes, G. H. L. G. A. (2021). A formal quantitative study of privacy in the publication of official educational censuses in Brazil. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

Organização das Nações Unidas (2014). Fundamental Principles of Official Statistics (A/RES/68/261 from 29 January 2014). Disponível em: `https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx`.

Pollard, T. J., Johnson, A. E., Raffa, J. D., Celi, L. A., Mark, R. G., and Badawi, O. (2018). The eicu collaborative research database, a freely available multi-center database for critical care research. *Scientific data*, 5(1):1–13.

Sarmin, F. J., Sarkar, A. R., Wang, Y., and Mohammed, N. (2024). Synthetic data: Revisiting the privacy-utility trade-off. *arXiv preprint arXiv:2407.07926*.

Smith, G. (2009). On the foundations of quantitative information flow. In *International Conference on Foundations of Software Science and Computational Structures*, pages 288–302. Springer.

Sweeney, L. (2000). Simple Demographics Often Identify People Uniquely. Disponível em: `https://kilthub.cmu.edu/articles/Simple_Demographics_Often_Identify_People_Uniquely/6625769/1`.