Segurança e Robustez de Estratégias para Anonimização e De-Anonimização de Dados Sensíveis de Localização

Marília M. Biz[1], **Henry C. Nunes**[2], **Bruno V. Lima**[1], **Weverton Cordeiro**[1]

[1] Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
[2] Escola Politécnica – Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

Abstract. *Releasing detailed spatial data benefits public health research but introduces concrete privacy risks, particularly in urban epidemiological surveillance. In this paper, we examine four families of anonymization techniques (generalization, microaggregation, permutation, and geographic differential privacy) applied to georeferenced Aedes aegypti surveillance data from a Porto Alegre. We compare these techniques using a unified experimental pipeline that jointly measures predictive utility (e.g., vector density forecasting) and robustness to re-identification. Evaluating across four attack models and three levels of auxiliary information leakage using paired tests with Holm-Bonferroni correction, we find that no anonymization variant causes a statistically significant loss of utility compared to the original data. Among the evaluated methods, microaggregation with* $k = 10$ *offers the best empirical balance: it maintains baseline predictive performance while providing robust protection against a temporal reconstruction attack that exposes a critical vulnerability in naive Geo-DP implementations that add independent noise to sequential observations.*

Resumo. *A disponibilização de dados espaciais detalhados beneficia a pesquisa em saúde pública, mas introduz riscos de privacidade concretos, particularmente na vigilância epidemiológica urbana. Neste artigo, examinamos quatro famílias de técnicas de anonimização (generalização, microagregação, permutação e privacidade diferencial geográfica) aplicadas a dados georreferenciados de vigilância do Aedes aegypti de Porto Alegre. Comparamos essas técnicas por meio de um pipeline experimental unificado que mede conjuntamente a utilidade preditiva (por exemplo, a previsão de densidade vetorial) e a robustez contra a reidentificação. Avaliando quatro modelos de ataque e três níveis de vazamento de informações auxiliares por meio de testes pareados com correção de Holm-Bonferroni, constatamos que nenhuma variante de anonimização causa uma perda de utilidade estatisticamente significativa em comparação aos dados originais. Entre os métodos avaliados, a microagregação com* $k = 10$ *oferece o melhor equilíbrio empírico: ela mantém o desempenho preditivo de referência (baseline) ao mesmo tempo em que fornece uma proteção robusta contra um ataque de reconstrução temporal que expõe uma vulnerabilidade crítica em implementações ingênuas de privacidade diferencial geográfica que adicionam ruído independente a observações sequenciais.*

1. Introdução

O crescimento exponencial de dispositivos IoT e da computação ubíqua gera volumes sem precedentes de dados espaciais, que podem ser extremamente úteis em pesquisas científicas, desenvolvimento de novos produtos e soluções, etc. Há diversos exemplos na literatura de como esses dados podem ser empregados em pesquisas científicas, incluindo detecção de comunidades [Khanfor et al. 2019], vigilância epidemiológica

[Aleixo et al. 2022], entre outros. Em muitos casos, no entanto, os dados são de natureza sensível (por ex., dados relativos à saúde ou ao comportamento). Assim, a exposição indevida desses dados pode representar riscos significativos à privacidade individual.

A reidentificação de indivíduos em conjuntos de dados espaciais, mesmo depois de remover nomes e outros identificadores diretos, deixou de ser hipótese teórica há tempos. Em bases de grande escala, quatro pontos espaço-temporais bastam para identificar a maioria das pessoas [de Montjoye et al. 2013]. Ainda assim, boa parte dos datasets públicos de saúde e vigilância continua sendo "anonimizada" com arredondamento grosseiro de coordenadas ou agregação *ad hoc*, sem garantias formais e quase sempre sem testes sistemáticos de robustez [GDPR Register 2020, Botcrawl Security Reports 2025], com muitos incidentes observados. Na prática, há uma lacuna entre o que a literatura de segurança mostra ser possível atacar e as soluções que são adotadas para se defender.

Em vigilância epidemiológica [Fraccaro et al. 2019], coordenadas de alta precisão de ocorrência de casos são essenciais para análises sobre a dinâmica dos fenômenos de interesse. Ao mesmo tempo, elas fornecem ao adversário uma superfície de ataque rica para reidentificação e uso malicioso das informações deanonimizadas, sob risco de prejuízos aos indivíduos que tiveram suas informações coletadas. Assim, instituições de saúde acabam tornando-se muito resistentes no compartilhamento de dados mesmo que anonimizados. Dado esse contexto, um grande desafio que se impõe é "*Como permitir acesso a grandes volumes de dados sensíveis para investigações científicas, garantindo que as pesquisas possam ser realizadas com a maior qualidade possível, e preservando a privacidade e inviolabilidade dos dados pessoais?*".

Diversas técnicas de anonimização tem sido utilizadas para proteger a privacidade de dados, desde generalização e microagregação (k-anonimato) [Sweeney 2002, Domingo-Ferrer and Mateo-Sanz 2002, Machanavajjhala et al. 2007] a permutações espaciais e privacidade diferencial geográfica [Andrés et al. 2013, Dwork et al. 2006]. Como cada abordagem possui garantias e custos distintos, a escolha prática é complexa. Em contrapartida, a literatura documenta ataques de desanonimização que usam informações auxiliares para reidentificar registros sem identificadores diretos. Destacam-se as estratégias de ligação direta e probabilística por proximidade espacial [Narayanan and Shmatikov 2008], além de ataques baseados em ambiguidade espacial e indistinguibilidade [Andrés et al. 2013, Shokri et al. 2012]. Estudos empíricos comprovam que a falta de uma avaliação rigorosa desses mecanismos pode comprometer a privacidade dos dados espaciais [Zang and Bolot 2011].

Apesar da significativa atenção que o tópico tem recebido na literatura, as propostas de técnicas de anonimização de dados avaliam de forma limitada o impacto da aplicação dessas técnicas nas investigações que são feitas *à posteriori* sobre os dados anonimizados. Em sua maioria, a literatura compreende estudos que agregam resultados prévios para comparar diferentes técnicas em termos de privacidade e utilidade dos dados. Assim, efeitos colaterais da anonimização, principalmente relacionados à perda da qualidade e representatividade dos dados, podem prejudicar investigações independentes e subsequentes (por exemplo, [Ribeiro et al. 2023]) que aplicam técnicas de aprendizado de máquina para previsão de séries temporais a partir de dados anonimizados. Assim, no presente artigo nos concentramos nas seguintes questões de pesquisa:

- Qual o impacto das técnicas de anonimização de dados de geolocalização nas soluções baseadas em aprendizado de máquina para predição de séries e tendências?
- Qual a resiliência das técnicas de anonimização de dados de geolocalização diante de ataques distintos de desanonimização?

Este trabalho apresenta uma avaliação empírica e sistemática de quatro famílias de técnicas de anonimização espacial sob quatro modelos de ataque de reidentificação. Consideramos um adversário com acesso aos dados publicados e conhecimento auxiliar parcial dos registros originais, em proporções compatíveis com vazamentos documentados no setor de saúde no Brasil [GDPR Register 2020, Botcrawl Security Reports 2025]. Nossas contribuições estendem-se em três frentes:

- **Framework Experimental:** Disponibilizamos um arcabouço reprodutível[1] que unifica a avaliação de utilidade preditiva e robustez adversarial de dados espaciais.
- **Avaliação Empírica Abrangente:** Avaliamos quatro modelos de ataque (ligação direta por vizinhança, ligação probabilística Top-k, ambiguidade espacial e reconstrução temporal) sobre dados epidemiológicos reais.
- **Análise de Sensibilidade e Implicações Práticas:** Demonstramos, por meio de análise estatística rigorosa, que conclusões consolidadas da literatura comparativa podem se inverter ao variar hiperparâmetros críticos (ε, k e precisão decimal), impactando diretamente a escolha de mecanismos em cenários sensíveis.

O restante do artigo está organizado da seguinte forma. A Seção 2 posiciona o trabalho frente à literatura. A Seção 3 descreve materiais e métodos. A Seção 4 apresenta as técnicas de anonimização. A Seção 5 avalia utilidade preditiva. A Seção 6 formaliza o modelo de adversário e os ataques. A Seção 7 apresenta a avaliação experimental de robustez, enquanto que a Seção 8 conclui o artigo.

2. Trabalhos Relacionados

A literatura em privacidade de dados espaciais evoluiu em duas frentes complementares: mecanismos defensivos de anonimização e vetores ofensivos de reidentificação. *Surveys* consolidados [Krumm 2009, Primault et al. 2019] oferecem panoramas amplos. Esta seção discute alguns trabalhos mais relevantes e identifica a lacuna que motiva o estudo, a partir de levantamento conduzido na ACM DL, IEEE Xplore, Scopus e SBC OpenLib (2002–2025) e de busca reversa nos trabalhos seminais [Sweeney 2002, Andrés et al. 2013, de Montjoye et al. 2013].

Mecanismos de Anonimização Espacial. A primeira geração de técnicas, baseada em truncamento ou arredondamento de coordenadas, oferece proteção essencialmente por obscuridade. Estudos empíricos em cenários de *wearables* [Park et al. 2021] demonstram a insuficiência dessas abordagens: a reidentificação permanece possível via metadados auxiliares. O k-anonimato [Sweeney 2002] foi formalizado como a exigência de que cada registro seja indistinguível de pelo menos $k - 1$ outros. Em dados espaciais numéricos, a microagregação [Domingo-Ferrer and Mateo-Sanz 2002] é a implementação mais comum, agrupando registros próximos e substituindo coordenadas pelo centroide do grupo. Limitações conhecidas (ataques de homogeneidade) motivaram extensões como ℓ-diversidade [Machanavajjhala et al. 2007]. No contexto brasileiro, trabalhos recentes [Coelho et al. 2024, Coelho et al. 2025] propõem métodos para seleção automatizada de k, e heurísticas bio-inspiradas para agrupamento são exploradas em [Pimenta et al. 2025]. Esses trabalhos focam no aprimoramento da técnica defensiva sem avaliação exaustiva de robustez frente a ataques específicos.

A noção de *geo-indistinguishability* [Andrés et al. 2013] estendeu privacidade diferencial [Dwork et al. 2006] para o domínio espacial, com a garantia formal de que dois

[1] https://github.com/mariliamatosbiz/article-spatial-anonymization-robustness

pontos próximos sejam indistinguíveis até um fator $e^{\varepsilon d}$. A literatura ainda discute intensamente a escolha apropriada de ε em cenários reais [Turgay et al. 2023].

Ataques de Reidentificação. O caso Netflix [Narayanan and Shmatikov 2008] mostrou que mesmo *datasets* grandes e esparsos são vulneráveis quando o adversário possui informação auxiliar parcial. Esse resultado foi estendido para dados de mobilidade [de Montjoye et al. 2013], com observação de unicidade extrema em *call detail records*. O ataque ótimo de um adversário contra mecanismos de proteção de localização foi formalizado em [Shokri et al. 2012], estabelecendo o erro esperado como métrica natural de privacidade. Outros trabalhos [Zang and Bolot 2011] concluem que anonimização ingênua é fundamentalmente insuficiente em populações urbanas reais. Uma revisão sistemática anterior [El Emam et al. 2011] identificou que estudos raramente avaliam conjuntamente múltiplos vetores, observação que motiva diretamente este trabalho.

Estudos Integrados e Lacuna Atual. Poucos trabalhos avaliam, conjuntamente os mecanismos defensivos e ataques sobre dados espaciais reais. O metodologicamente mais próximo [Jin et al. 2023] concentra-se em trajetórias completas, o cenário de pontos discretos fica fora de seu escopo. Estudos em *wearables* [Park et al. 2021] restringem-se a anonimização ingênua sem mecanismos formais, e há também testes de k-anonimato em *GPS traces* [Hoh et al. 2007], em escala restrita. As demais referências [Lin et al. 2021, Sampaio et al. 2023, Majeed and Lee 2021, Löbner et al. 2021]são análises secundárias sem experimentos primários comparativos.

Embora possua investigações maduras em cada eixo isolado, não encontramos, no nosso levantamento da literatura, avaliação empírica que combine vários paradigmas de anonimização espacial, múltiplos vetores de ataque, dados reais de um domínio válido e relevante e análise estatística com intervalos de confiança e sensibilidade.

3. Materiais e Métodos

3.1. Natureza dos Dados

Os dados provêm do projeto público, mantido pela Secretaria Municipal de Saúde de Porto Alegre no estado do Rio Grande do Sul. O *dataset* contém registros de inspeções a armadilhas entomológicas instaladas em logradouros públicos, totalizando aproximadamente 235 mil inspeções entre 2019 e 2023 após filtros de qualidade. A integração com dados meteorológicos do INMET é parte essencial do *pipeline* preditivo (Seção 5). Como a série meteorológica disponível, inicia em janeiro de 2019, o estudo adota como período os anos de 2019 a 2023, garantindo correspondência temporal entre variáveis entomológicas e climáticas.

É importante destacar que cada registro corresponde a uma observação de armadilha, não de pessoa, pois as coordenadas identificam o local físico do equipamento, não cidadãos. A relevância da anonimização decorre do fato de que coordenadas precisas podem revelar, por inferência, imóveis privados, servindo de banco de prova para*datasets* análogos com dados pessoais.

3.2. Filtros de Qualidade

Realizada a aplicação de filtros em cascata para tornar os registros geograficamente válidos, sendo eles, a remoção de coordenadas nulas, sinais geográficos inválidos, com tolerância de aproximadamente 10 km dos limites administrativos do município. Para o período 2019 à 2023, o conjunto contém $\sim$224 mil registros geograficamente válidos. Em etapa subsequente do *pipeline* preditivo, registros com inspeção ausente são descartados, pois a agregação semanal por micro-região exige carimbo temporal válido.

3.3. Pipeline Experimental

A avaliação compreende três etapas integradas em um único *pipeline* reproduzível: (i) anonimização espacial sobre as coordenadas brutas, produzindo 10 variantes parametrizadas (Tabela 1); (ii) avaliação de utilidade preditiva sob validação cruzada temporal e múltiplos classificadores (Seção 5); e (iii) avaliação de privacidade sob quatro modelos de ataque com diferentes níveis de conhecimento auxiliar (Seção 6).

Tabela 1. Hiperparâmetros das técnicas, ataques e *pipeline* preditivo.

Componente	Parâmetro	Valor(es) avaliado(s)	Fonte / Justificativa
Anonimização			
Generalização	casas decimais	$\{2\}^{\dagger}$ (precisão $\sim$ 1,1 km)	[Sweeney 2002, Murthy et al. 2019]
Microagregação	k	$\{2, 5, 10\}$	[Domingo-Ferrer and Mateo-Sanz 2002, Machanavajjhala et al. 2007]; k=10 valor canônico
Permutação		embaralhamento aleatório de lat./lon.	[Majeed and Lee 2021]
Privacidade Diferencial	ε (1/km)	$\{0{,}1;\ 0{,}5;\ 1{,}0;\ 2{,}0;\ 5{,}0\}$	[Andrés et al. 2013, Turgay et al. 2023]; varredura para mapear *trade-off*
Ataques de reidentificação			
A1 (NN-Linkage)	raio (m)	$\{50, 100, 200\}$	[Shokri et al. 2012]; precisão GPS urbano
A2 (Top-k)	k	$\{3, 5, 10\}$	[Narayanan and Shmatikov 2008]; ligação *Top-k*
A3 (Ambiguidade)	raio (m)	$\{50, 100, 200, 500\}$; 200 m principal	[Shokri et al. 2012]; quarteirão urbano
A4 (Reconstrução temp.)	janela	7 dias; mediana espacial	[Jin et al. 2023]; ciclo de inspeção
Modelo de adversário e pipeline *preditivo*			
Adversário	vazamento aux.	$\{1\%, 5\%, 10\%\}$	[Narayanan and Shmatikov 2008]; vazamentos parciais documentados
Validação temporal	estratégia	`TimeSeriesSplit` (5 *folds* expansivos)	prática consolidada em séries temporais
Múltiplas execuções	*seeds*	10 (técnicas estocásticas)	ICs *bootstrap* 95% estáveis
Inferência estatística	IC + teste	*Bootstrap* 95% + Wilcoxon + Holm-Bonferroni	[El Emam et al. 2011]
Classificadores	comparação	XGBoost, Random Forest, Reg. Logística	cobertura: *boosting*, *bagging*, linear
Limiar classificação	otimização	maximização de F_1 no treino do *fold*	[Saito and Rehmsmeier 2015]

† A variante `dec1` (1 casa decimal, $\sim$ 11 km) foi inicialmente avaliada e excluída por colapso espacial: reduz a cidade a 8 pontos distintos, insuficiente para clusterização em 15 micro-regiões.

Adotado `TimeSeriesSplit` com cinco *folds* expansivos sobre o período 2019 a junho de 2023. Cada *fold* expande a janela de treino e desloca a janela de teste no tempo, respeitando a ordem cronológica, propriedade essencial para evitar vazamento temporal. A extração dos dados publicados pela cidade de Porto Alegre contém registros até 09 de junho de 2023, e o uso de `TimeSeriesSplit` permite aproveitar toda a série disponível, ampliando a amostragem efetiva de avaliação. O percentil 80 utilizado para definir a variável-alvo Y_{surto} é calculado dentro de cada *fold*, exclusivamente sobre os dados de treino correspondentes, evitando vazamento de informação.

3.4. Reprodutibilidade e Considerações Éticas

O *framework* experimental completo (código, scripts, modelos, *seeds*, resultados intermediários) está publicamente disponível [2]. Cada experimento é reproduzível em ambiente Google Colab. Conforme a Resolução CNS nº 510/2016, Art. 1º, §1º, pesquisas com informações de domínio público são dispensadas de registro no sistema CEP/CONEP, hipótese aplicável ao *dataset* aqui utilizado. Quanto à LGPD, o Art. 5º, inciso I, define dado pessoal como "informação relacionada a pessoa natural identificada ou identificável", onde as coordenadas empregadas correspondem a localizações de armadilhas instaladas em logradouros públicos, não constituindo dados pessoais no sentido legal.

4. Técnicas de Anonimização

Cada família descrita representa um paradigma distinto: a generalização é sintática e determinística, a microagregação é sintática, com particionamento aleató-

[2]https://github.com/mariliamatosbiz/article-spatial-anonymization-robustness

rio, mas produz mapeamento fixo por ponto, e atinge k-anonimato pela substituição de coordenadas por centroides de grupos, permutação é estocástica e preserva distribuições marginais, mas desfaz a associação espacial original, privacidade diferencial geográfica é o único mecanismo aqui com garantia formal parametrizável. Mecanismos derivados como ℓ-diversidade [Machanavajjhala et al. 2007] ou MDAV [Domingo-Ferrer and Mateo-Sanz 2002] são variantes algorítmicas desses paradigmas e ficam como objeto de comparação direta em trabalhos futuros (Seção 8).

Seja $\mathcal{D} = \{r_1, \ldots, r_n\}$ o conjunto de registros originais, com $r_i = (lat_i, lon_i, t_i, a_i)$. Cada técnica define uma transformação $\mathcal{M} : \mathcal{D} \to \mathcal{D}^*$ que preserva atributos não-espaciais a_i, modificando apenas as coordenadas.

4.1. Generalização

A generalização reduz a precisão por arredondamento decimal [Sweeney 2002, Murthy et al. 2019], seja $d \in N$ o número de casas decimais preservadas:

$$\mathcal{M}^{(d)}_{\text{gen}}(lat_i, lon_i) = (\text{round}(lat_i, d),\ \text{round}(lon_i, d)). \tag{1}$$

No extremo sul do Brasil ($lat \approx -30°$), $d = 1$ corresponde a ~ 11 km, $d = 2$ a $\sim 1{,}1$ km. A técnica é determinística. Avaliamos $d = 2$ (a variante $d = 1$ é excluída por colapso espacial, conforme nota da Tabela 1).

4.2. Microagregação (Implementação de *k*-Anonimato)

É comum a confusão entre k-anonimato e microagregação. k-anonimato [Sweeney 2002] é um modelo de privacidade, que $\mathcal{D}^*$ satisfaz k-anonimato se cada registro é indistinguível de pelo menos outros $k - 1$. Microagregação [Domingo-Ferrer and Mateo-Sanz 2002] é o *algoritmo* que atinge k-anonimato em dados numéricos. Dados $k \geq 2$:

1. **Particionamento:** construir partição $\mathcal{P} = \{G_1, \ldots, G_m\}$ de $\mathcal{D}$ com $|G_j| \geq k\ \forall j$;
2. **Agregação:** para cada $r_i \in G_j$, definir $\mathcal{M}^{(k)}_{\text{micro}}(lat_i, lon_i) = (\bar{lat}_j, \bar{lon}_j)$, onde $(\bar{lat}_j, \bar{lon}_j)$ é o centroide geográfico de G_j.

Adotamos particionamento aleatório uniforme, controlado por *seed* de reprodutibilidade. Avaliamos $k \in \{2, 5, 10\}$.

4.3. Permutação

Permutação modifica coordenadas mediante embaralhamento aleatório independente de latitude e longitude, preservando distribuições marginais mas desfazendo a associação espacial original, sejam $\pi_{lat}, \pi_{lon} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ permutações aleatórias independentes:

$$\mathcal{M}_{\text{perm}}(lat_i, lon_i) = (lat_{\pi_{lat}(i)},\ lon_{\pi_{lon}(i)}). \tag{2}$$

A independência é crítica: quebra a correlação espacial. A técnica é estocástica.

4.4. Privacidade Diferencial Geográfica

O paradigma mais forte avaliado é a *geo-indistinguishability* [Andrés et al. 2013], extensão de privacidade diferencial [Dwork et al. 2006]:

Definição 1 (ε-Geo-Indistinguishability)**.** *Um mecanismo $\mathcal{M}$ satisfaz ε-geo-indistinguishability se, para quaisquer $x, x' \in R^2$ e qualquer mensurável $Z \subseteq R^2$:*

$$\Pr[\mathcal{M}(x) \in Z] \leq e^{\varepsilon \cdot d(x,x')} \cdot \Pr[\mathcal{M}(x') \in Z], \tag{3}$$

em que $d(x, x')$ é a distância em km.

O parâmetro ε tem unidade km^{-1} e atua como orçamento de privacidade, a realização canônica é o mecanismo Planar Laplace, gera-se deslocamento polar (r, θ) com $\theta \sim \text{Uniform}(0, 2\pi)$ e $r \sim \text{Gamma}(k{=}2, \text{scale}{=}1/\varepsilon)$ (com r em km), e aplica-se o deslocamento $(r\cos\theta, r\sin\theta)$, após conversão de metros para graus geográficos, ao ponto verdadeiro. O raio mediano do deslocamento é aproximadamente $1{,}68/\varepsilon$ km, validado empiricamente: $\varepsilon = 0{,}1$ produz mediana de $\sim$16,9 km; $\varepsilon = 1{,}0$, $\sim$1,68 km; $\varepsilon = 5{,}0$, $\sim$340 m. A varredura $\varepsilon \in \{0{,}1;\ 0{,}5;\ 1{,}0;\ 2{,}0;\ 5{,}0\}$ permite mapear explicitamente o *trade-off* [Turgay et al. 2023].

A Figura 1 apresenta a distribuição dos pontos no dataset, incluindo a distribuição original e a distribuição resultante após processos de anonimização. A Figura 2 detalha as distribuições individuais entre o ponto original e o resultante anonimizado por técnica de anonimização, evidenciando as diferenças de forma, cauda longa em DP $\varepsilon{=}0{,}1$, picos discretos em generalização determinística, distribuições com cauda em microagregação.

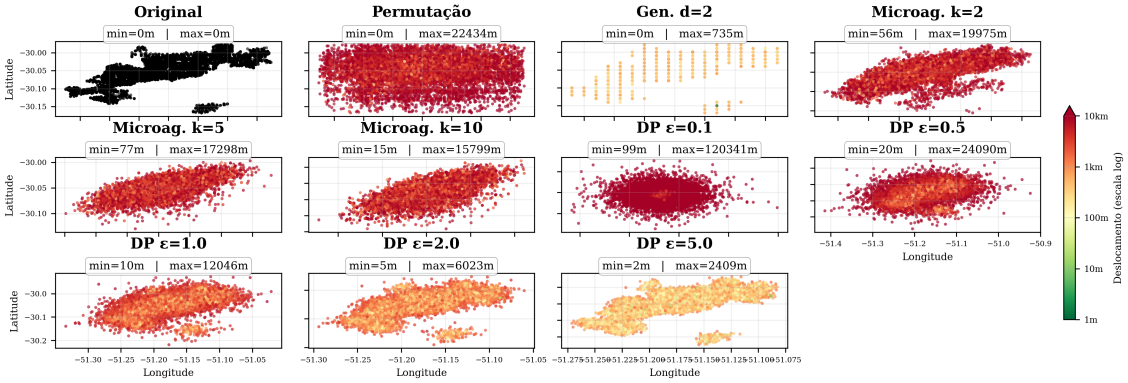

Figura 1. Distribuição espacial dos pontos anonimizados por técnica.

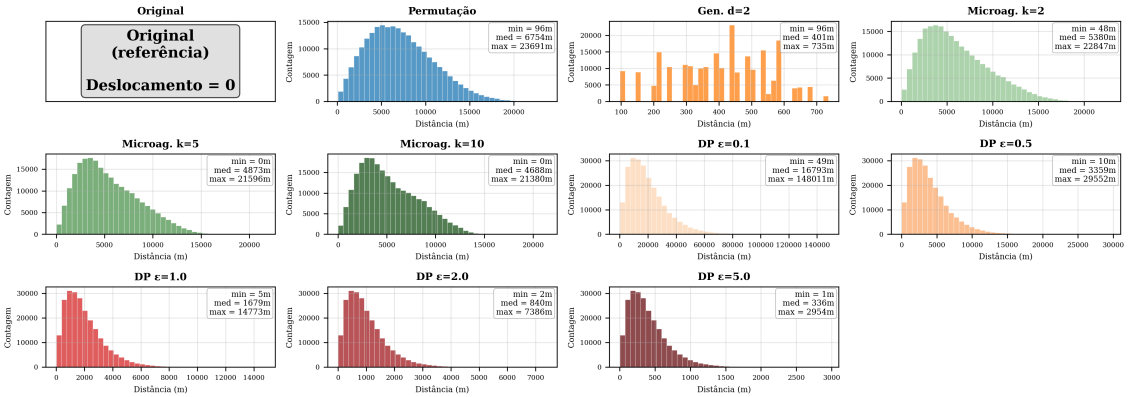

Figura 2. Histogramas do deslocamento em metros de cada ponto, por técnica.

A Figura 3, por sua vez, resume a distribuição empírica de deslocamentos para as 10 técnicas avaliadas sobre o *dataset*. Pode-se observar que generalização $d{=}2$ produz deslocamentos concentrados (mediana $\sim$ 400 m), microagregação e permutação distribuem-se em $\sim$ 5 km, e DP segue a relação prevista entre ε e magnitude do ruído.

4.5. Complexidade Computacional e *Overhead*

As técnicas avaliadas têm custo computacional baixo. Generalização, permutação, microagregação aleatória e privacidade diferencial executam em tempo linear no número de registros n, com custo constante por registro. Variantes que minimizam perda de informação, como MDAV [Domingo-Ferrer and Mateo-Sanz 2002], podem chegar a $O(n^2)$ no pior caso, mas não foram avaliadas aqui. Empiricamente, sobre os $\sim$224 mil registros

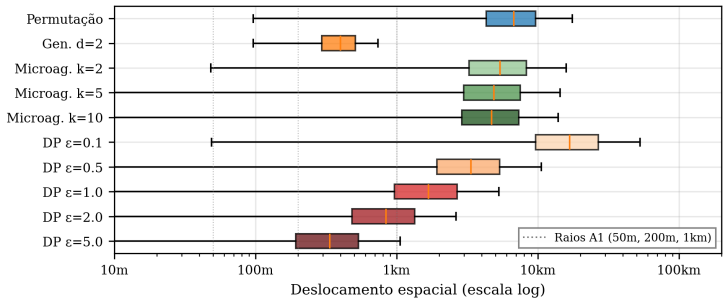

Figura 3. Distribuição do deslocamento espacial por técnica (escala log). Linhas pontilhadas indicam os raios operacionais $\{50, 200, 1\,000\}$ m utilizados na avaliação de ataques.

válidos, qualquer técnica é aplicada rapidamente no Google Colab, o que é desprezível frente ao custo das etapas posteriores, treinamento dos modelos preditivos ($\sim$67 min para 1 380 *fits*) e execução dos ataques A1–A4 ($\sim$46 min). Para o volume típico de *datasets* municipais, a anonimização não é gargalo em *pipelines offline* de publicação.

5. Modelo Preditivo e Avaliação de Utilidade

5.1. Tarefa de Predição

A tarefa adotada é a predição binária de surto semanal de *Aedes aegypti* em escala municipal, para cada semana t, o modelo prevê $Y_t \in \{0, 1\}$:

$$Y_t = \begin{cases} 1 & \text{se } X_t^{\text{total}} > \tau_{P80}^{\text{treino}} \\ 0 & \text{caso contrário,} \end{cases} \tag{4}$$

onde X_t^{total} é o total semanal de *Aedes aegypti* coletados na cidade, e $\tau_{P80}^{\text{treino}}$ é o percentil 80 calculado exclusivamente sobre o treino do fold corrente, evitando vazamento temporal onde o percentil 80 reflete prática consolidada em vigilância, sendo que os surtos correspondem aos eventos do quintil superior da distribuição.

5.2. Engenharia de *Features*

A construção segue os seguintes princípios: **Antecipação por defasagem.** Para predizer Y_t, utilizamos apenas observações de semanas anteriores ($t-2$, $t-3$, $t-4$). A defasagem mínima de duas semanas garante horizonte operacionalmente útil e elimina respostas fáceis por autocorrelação. **Decomposição espacial.** Coordenadas anonimizadas são agrupadas em $N = 15$ micro-regiões via `K-Means` (`random_state=42`). Para cada região r e semana t, computamos $X_{t,r}$, o total de *Aedes aegypti* coletados. **Variáveis climáticas com base biológica.** Quatro *features* climáticas semanais (chuva_acum_21d, temp_min_media_14d, dias_temp_min_20c, dias_sem_chuva), defasadas nos mesmos *lags*. Total: $(15 + 1) \times 3 + 4 \times 3 = 60$ *features*. Em resumo, o vetor de entrada combina 16 atributos entomológicos derivados da cidade de Porto Alegre (totais semanais por micro-região, $X_{t,r}$, e total municipal X_t^{total}) com 4 atributos climáticos do INMET (`chuva_acum_21d`, `temp_min_media_14d`, `dias_temp_min_20c` e `dias_sem_chuva`), todos defasados em três janelas. O alvo $Y_t \in \{0, 1\}$ é derivado do total municipal contra o percentil 80 do treino do fold.

5.3. Classificadores e Treinamento

Foi realizada a avaliação dos seguintes classificadores: **XGBoost** (modelo principal, com `scale_pos_weight`), **Random Forest** (`class_weight=balanced`) e **Regressão**

Logística (L_2, balanceada). Para cada combinação técnica $\times$ *seed* $\times$ *fold* $\times$ classificador, o *threshold* é otimizado sobre o treino para maximizar F_1 [Saito and Rehmsmeier 2015]. Essa estratégia evita o viés do *threshold* padrão $0{,}5$, inadequado para classes desbalanceadas, sem usar dados de teste.

5.4. Resultados de Utilidade

A Tabela 2 apresenta as métricas de desempenho para cada técnica, com intervalos de confiança 95% via *bootstrap* (1 000 amostras). Valores são medianas sobre 10 *seeds* $\times$ 5 *folds* $\times$ 3 classificadores ou sobre 5 *folds* $\times$ 3 classificadores, a última coluna apresenta o p-valor de Wilcoxon pareado vs `original`, ajustado por Holm-Bonferroni.

Tabela 2. Avaliação de utilidade preditiva. Variação total: ΔAUC $= 0{,}049$ pontos ($5{,}4\%$). Nenhuma técnica rejeita H_0 a $\alpha = 0{,}05$.

Técnica	AUC	Recall	F1	Bal. Acc.	p Holm
`original`	0,918 [0,80–0,97]	0,714	0,727	0,749	
`microagregacao_k10`	0,904 [0,88–0,93]	0,710	0,652	0,727	1,00
`permutacao`	0,885 [0,86–0,92]	0,699	0,656	0,720	1,00
`microagregacao_k5`	0,885 [0,86–0,92]	0,710	0,632	0,720	1,00
`dp_eps_5.0`	0,879 [0,86–0,92]	0,647	0,632	0,720	0,94
`dp_eps_0.5`	0,875 [0,85–0,91]	0,706	0,637	0,694	1,00
`microagregacao_k2`	0,873 [0,85–0,90]	0,647	0,600	0,692	1,00
`dp_eps_2.0`	0,871 [0,85–0,90]	0,602	0,596	0,696	1,00
`generalizacao_dec2`	0,871 [0,83–0,94]	0,765	0,700	0,786	1,00
`dp_eps_1.0`	0,871 [0,86–0,92]	0,710	0,632	0,717	1,00
`dp_eps_0.1`	0,869 [0,85–0,91]	0,710	0,638	0,714	1,00

5.5. Análise Estatística e Discussão

A análise adotou bootstrap não-paramétrico para construção dos IC 95% (1 000 reamostragens estratificadas). Aplicou-se o teste pareado de Wilcoxon contra o *baseline* original, pareando por fold e classificador, com mediana entre seeds; o teste foi preferido ao t-pareado pela robustez à não-normalidade. A correção de Holm-Bonferroni foi aplicada sobre 10 comparações simultâneas por métrica.

Os resultados levam a três observações. Primeiro, a variação inter-técnica é estreita: o AUC mediano vai de 0,869 (`dp_eps_0.1`) a 0,918 (`original`), amplitude de 0,049 pontos ou 5,4%. Como `dp_eps_0.1` corresponde a deslocamentos medianos da ordem de 17 km, superiores à extensão do próprio município, a preservação de utilidade chama atenção. Segundo, essa proximidade não é apenas visual: após a correção de Holm-Bonferroni, todas as comparações contra o *baseline* retornam $p_{\text{Holm}} \geq 0{,}94$ para AUC, padrão que se repete nas demais métricas. Terceiro, a explicação é arquitetural: ao agregar inspeções por micro-regiões e janelas semanais, o pipeline produz *features* cujo valor depende da distribuição coletiva, e não das posições exatas, de modo que o ruído individual da anonimização é diluído nessa agregação. Esse efeito é visível em `microagregacao_k10` (AUC 0,904, segundo melhor), em que a agregação em grupos de dez é compatível com a agregação semanal por região.

6. Modelo e Estratégias de Ataque

6.1. Modelo de Ataque

Adotamos um modelo de ataque inspirado em trabalhos consolidados da área [Narayanan and Shmatikov 2008, Andrés et al. 2013, Shokri et al. 2012]. Seja $\mathcal{A}$ um adversário polinomialmente limitado. Assume-se que $\mathcal{A}$ tem acesso integral ao *dataset* anonimizado $\mathcal{D}^*$ e dispõe de conhecimento auxiliar parcial dos registros originais, $\mathcal{D}_{\text{aux}} \subset \mathcal{D}$, em proporção $|\mathcal{D}_{\text{aux}}|/|\mathcal{D}| \in \{1\%, 5\%, 10\%\}$, valores compatíveis com vazamentos parciais. Em termos computacionais, $\mathcal{A}$ pode calcular distâncias *haversine*

(distâncias entre dois pontos de uma esfera a partir de suas latitudes e longitudes), realizar consultas espaciais e treinar modelos sobre $\mathcal{D}^*$.

Por outro lado, o adversário não tem acesso ao conjunto $\mathcal{D} \setminus \mathcal{D}_{\text{aux}}$ nem a identificadores diretos; desconhece o mecanismo $\mathcal{M}$ e seus parâmetros (configuração *black-box*); e não dispõe de canais laterais. Seu objetivo, para cada $r_{\text{aux}} \in \mathcal{D}_{\text{aux}}$, é determinar qual registro $r^* \in \mathcal{D}^*$ corresponde a r_{aux}.

6.2. Estratégias de Ataque

A1 - Reidentificação Direta por Vizinhança Espacial. Dado r_{aux} e raio R, o ataque é bem-sucedido se $d_{\text{hav}}(r_{\text{aux}}, \mathcal{M}(r_{\text{aux}})) \leq R$, onde $\mathcal{M}(r_{\text{aux}}) \in \mathcal{D}^*$ é o registro anonimizado correspondente. A taxa de sucesso é:

$$\tau_{\text{A1}}(R) = \frac{1}{|\mathcal{D}_{\text{aux}}|} \sum_{r \in \mathcal{D}_{\text{aux}}} \mathbf{1}[d_{\text{hav}}(r, \mathcal{M}(r)) \leq R]. \tag{5}$$

Avaliamos $R \in \{50, 100, 200\}$ m (200 m como referência principal).

A2 - Ligação Probabilística *Top-k*. Generaliza A1 ao considerar os k registros de $\mathcal{D}^*$ mais próximos. Seja $\text{NN}_k(r_{\text{aux}}; \mathcal{D}^*)$ esse conjunto:

$$\tau_{\text{A2}}(k) = \frac{1}{|\mathcal{D}_{\text{aux}}|} \sum_{r \in \mathcal{D}_{\text{aux}}} \mathbf{1}[\mathcal{M}(r) \in \text{NN}_k(r; \mathcal{D}^*)]. \tag{6}$$

Avaliamos $k \in \{3, 5, 10\}$.

A3 - Anonimato Efetivo por Ambiguidade Espacial. Inverte a perspectiva: mede o tamanho do conjunto de candidatos plausíveis em torno de cada $r^* \in \mathcal{D}^*$. Para cada r^* e raio R, o conjunto de candidatos é $C_R(r^*) = \{r \in \mathcal{D} : d_{\text{hav}}(r, r^*) \leq R\}$, e a ambiguidade média é $\bar{a}(R) = \frac{1}{|\mathcal{D}^*|} \sum_{r^* \in \mathcal{D}^*} |C_R(r^*)|$. Valores maiores indicam maior proteção estrutural [Shokri et al. 2012]. O raio principal $R = 200$ m corresponde à escala típica de quarteirão urbano da cidade de Porto Alegre e à precisão usual de GPS de *smartphone* em zona densa; os demais raios (50, 100 e 500 m) compõem análise de sensibilidade. Para acelerar a consulta, usamos `BallTree` (*haversine*) sobre subamostra aleatória de $5\,000$ pontos de $\mathcal{D}^*$ (*seed* fixa), sem perda perceptível na média $\bar{a}(R)$.

A4 - Reconstrução Temporal por Agregação. Explora a consistência temporal de observações anonimizadas. Se uma armadilha é inspecionada múltiplas vezes, mecanismos como permutação ou DP produzem $\mathcal{M}(r_t)$ distintos a cada inspeção, mas a mediana espacial das observações tende a convergir para a posição verdadeira [Jin et al. 2023]. Seja $T(r_{\text{aux}})$ o conjunto de observações anonimizadas do mesmo ponto verdadeiro r_{aux}, dentro de janela Δt:

$$\hat{r}_{\text{aux}} = \text{mediana espacial}(T(r_{\text{aux}})), \tag{7}$$

e o ataque é bem-sucedido se $d_{\text{hav}}(\hat{r}_{\text{aux}}, r_{\text{aux}}) \leq R_{A4}$. Adotamos a mediana (não a média) por robustez a *outliers* da cauda Gamma do Planar Laplace. Avaliamos $\Delta t = 7$ dias.

7. Avaliação Experimental de Robustez Adversarial

A avaliação empírica dos quatro modelos de ataque é conduzida sobre as 10 técnicas de anonimização. Cada técnica estocástica é avaliada em 10 *seeds*, enquanto as determinísticas são executadas uma única vez. As tabelas desta seção trazem resultados para

$|\mathcal{D}_{\text{aux}}| = 10\%$; os cenários com 1% e 5% produzem resultados qualitativamente semelhantes e estão disponíveis no repositório.

7.1. A1 e A2 – Ligação Direta e *Top-k*

Sete das dez técnicas oferecem proteção superior a 98% contra reidentificação direta a $200\,\text{m}$. Observando a Tabela 3, pode perceber a degradação em DP é monotônica: $\tau_{A1}(200\,\text{m})$ cresce de $0{,}02\%$ (ε=0,1) para $26{,}39\%$ (ε=5,0), caracterizando este último como ponto operacional inadequado. `generalizacao_dec2` apresenta vulnerabilidade intermediária ($11{,}07\%$) devido à coincidência de pontos sobre a mesma célula de grade. O ataque A2 é quase totalmente neutralizado por técnicas baseadas em agregação (taxa 0% mesmo para k=10): agregações produzem deslocamentos da ordem de quilômetros, posicionando $\mathcal{M}(r_{\text{aux}})$ fora dos k candidatos mais próximos em uma cidade densamente coberta.

Tabela 3. Taxa de sucesso (%) dos ataques A1 e A2 por técnica, $|\mathcal{D}_{\text{aux}}| = 10\%$. Medianas sobre 10 *seeds*.

	A1 (NN-Linkage)			A2 (Top-k)		
Técnica	**50 m**	**100 m**	**200 m**	k=3	k=5	k=10
`permutacao`	0,00	0,01	0,05	0,00	0,00	0,00
`microagregacao_k2`	0,01	0,03	0,11	0,00	0,00	0,00
`microagregacao_k5`	0,01	0,04	0,13	0,00	0,00	0,00
`microagregacao_k10`	0,00	0,03	0,13	0,00	0,00	0,00
`dp_eps_0.1`	0,00	0,01	0,02	0,00	0,00	0,01
`dp_eps_0.5`	0,03	0,12	0,46	0,01	0,02	0,04
`dp_eps_1.0`	0,12	0,46	1,78	0,03	0,06	0,11
`dp_eps_2.0`	0,46	1,78	6,13	0,12	0,20	0,38
`generalizacao_dec2`	1,09	3,05	11,07	0,15	0,25	0,54
`dp_eps_5.0`	2,66	8,97	**26,39**	0,56	0,89	1,69

7.2. A3 - Ambiguidade Espacial

A Tabela 4 apresenta a ambiguidade média $\bar{a}(R)$ por técnica e raio. Microagregação com $k \geq 5$ sistematicamente supera as demais nos raios operacionais de 200–500 m. O valor baixo de `dp_eps_0.1` decorre do deslocamento para regiões de baixa densidade urbana, comportamento que não implica menor privacidade, pois a posição original não é revelada.

Tabela 4. Ambiguidade espacial média $\bar{a}(R)$ por técnica e raio

Técnica	**50 m**	**100 m**	**200 m**	**500 m**
`microagregacao_k10`	20,2	76,3	**312,6**	**1 950,6**
`microagregacao_k5`	21,4	78,3	301,4	1 898,5
`dp_eps_5.0`	18,7	72,5	282,9	1 694,0
`microagregacao_k2`	17,1	63,7	274,4	1 674,3
`generalizacao_dec2`	**26,6**	72,9	257,9	1 712,3
`dp_eps_2.0`	14,9	60,5	241,1	1 484,6
`dp_eps_1.0`	12,8	49,3	197,0	1 223,6
`permutacao`	15,4	43,1	149,9	1 001,8
`dp_eps_0.5`	9,0	35,6	140,5	878,1
`dp_eps_0.1`	1,5	6,3	25,5	158,1

7.3. A4 - Reconstrução Temporal

O ataque A4 expõe uma vulnerabilidade prática da DP geográfica quando aplicada sem coordenação entre observações do mesmo ponto físico. A Tabela 5 mostra que o erro mediano cai de $2\,092\,\text{m}$ (ε=0,1) para $42\,\text{m}$ (ε=5,0), queda de duas ordens de magnitude, enquanto técnicas de agregação mantêm erros acima de $4{,}6\,\text{km}$. Essa vulnerabilidade não invalida a garantia formal de ε-geo-indistinguishability, definida por observação e não por entidade ao longo do tempo; mitigações como *cache* de ruído por entidade ou composição sequencial do orçamento ε eliminam o ataque, mas exigem coordenação que *pipelines on-the-fly* tipicamente não implementam.

Tabela 5. Erro de reconstrução temporal $d_{\text{hav}}(\hat{r}, r_{\text{verdadeiro}})$ em metros, sob janela de 7 dias. Valores maiores indicam maior proteção.

Técnica	Mediana (m)	P_{95} (m)
`microagregacao_k10`	**4 718**	**10 240**
`permutacao`	4 695	10 689
`microagregacao_k5`	4 691	10 346
`microagregacao_k2`	4 682	10 507
`dp_eps_0.1`	2 092	7 066
`dp_eps_0.5`	418	1 413
`generalizacao_dec2`	401	627
`dp_eps_1.0`	209	707
`dp_eps_2.0`	105	353
`dp_eps_5.0`	**42**	**141**

A Figura 4 ilustra a queda abrupta da curva da DP, que cruza a zona crítica de 200 m entre ε=1,0 e ε=2,0, enquanto as linhas horizontais das técnicas determinísticas permanecem estáveis, reforçando sua robustez intrínseca ao ataque.

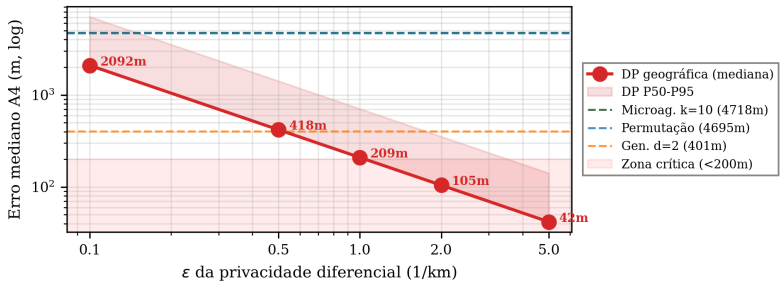

Figura 4. Colapso da proteção temporal da DP geográfica conforme ε cresce

7.4. Síntese dos Resultados

Os experimentos conduzidos permitem consolidar quatro observações que respondem diretamente às questões de pesquisa formuladas na Seção 1. A robustez da utilidade é propriedade do pipeline de agregação, não da anonimização, a preservação do AUC não decorre de as técnicas serem pouco intrusivas: `dp_eps_0.1` desloca pontos em mediana de $\sim$17 km, valor superior à própria extensão do município, e ainda assim mantém AUC de 0,869. O efeito decorre de três propriedades estruturais do pipeline: (i) a agregação espacial em $N = 15$ micro-regiões produz células cuja escala característica supera o deslocamento mediano da maioria das técnicas; (ii) a agregação temporal semanal dilui o ruído individual; e (iii) o alvo Y_t é definido sobre o total municipal X_t^{total}, invariante a qualquer redistribuição intramunicipal de coordenadas.

O ranqueamento entre técnicas é função do vetor de ataque adotado, nenhuma técnica domina sob todos os critérios: `dp_eps_5.0` figura entre as melhores em utilidade (AUC = 0,879) e entre as piores sob A1 (26,39% de reidentificação a 200 m) e A4 (42 m de erro). Avaliações restritas a um único vetor produziriam recomendações opostas sobre o mesmo mecanismo. A Garantia formal por observação não implica proteção por entidade, o ataque A4 reduz o erro de localização de 2 092 m (ε=0, 1) para 42 m (ε=5, 0) sem violar a definição de ε-geo-indistinguishability, enunciada por observação isolada e não sobre a sequência de observações de uma mesma entidade física. Trata-se de uma lacuna entre o modelo de ataque assumido na prova de privacidade e o modelo relevante em pipelines de publicação periódica. Mecanismos determinísticos, generalização, microagregação, produzem a mesma saída a cada reinspeção do mesmo ponto e não degradam sob agregação temporal. Mecanismos que sorteiam ruído de forma independente a cada observação expõem progressivamente a posição verdadeira, salvo se houver *cache* de ruído por entidade ou composição sequencial do orçamento ε.

7.5. Trade-off Privacidade-Utilidade

A microagregação com $k{=}10$ emerge como técnica de melhor compromisso multi-vetor, com proteção quase total contra A1/A2, máxima ambiguidade espacial nos raios operacionais, máxima resistência à reconstrução temporal e a maior utilidade preditiva entre as anonimizadas. A privacidade diferencial geográfica, apesar da garantia formal controlável via ε, apresenta vulnerabilidade não-trivial sob A4. Para aplicações que exigem garantia formal, recomendamos $\varepsilon \in [0{,}5; 1{,}0]$ como ponto operacional viável (proteção A1 $\geq 98{,}2\%$, erro A4 ≥ 209 m); valores $\varepsilon \geq 2{,}0$ não são recomendados ante a rápida deterioração temporal, e $\varepsilon{=}5{,}0$ mostra o limite onde a proteção colapsa. A Tabela 6 integra os resultados de utilidade e ataques.

Tabela 6. Posicionamento das 10 técnicas no plano privacidade-utilidade.

Técnica	Privacidade	Utilidade	Erro A4 (m)	Rank
`microagregacao_k10`	0,999	0,904	4 718	**2,0**
`microagregacao_k5`	0,999	0,885	4 691	2,5
`permutacao`	0,999	0,885	4 695	2,5
`microagregacao_k2`	0,999	0,873	4 682	4,5
`dp_eps_0.1`	1,000	0,869	2 092	5,5
`dp_eps_0.5`	0,995	0,875	418	5,5
`dp_eps_5.0`	0,736	0,879	42	7,0
`dp_eps_1.0`	0,982	0,871	209	7,5
`dp_eps_2.0`	0,939	0,871	105	8,0
`generalizacao_dec2`	0,889	0,871	401	8,5

8. Considerações Finais

Este artigo investigou o *trade-off* entre privacidade e utilidade em anonimização espacial sobre dados de vigilância entomológica de *Aedes aegypti*, avaliando quatro famílias de técnicas sob múltiplos vetores de ataque. A utilidade preditiva mostrou-se estruturalmente robusta à anonimização individual, com o AUC mediano em faixa estreita ($\Delta\text{AUC} = 0{,}049$ pontos) e sem diferenças significativas após a correção de Holm-Bonferroni ($p_{\text{Holm}} \geq 0{,}94$). A microagregação com $k{=}10$ ofereceu o melhor compromisso, dominando os quatro ataques avaliados. Já implementações de DP geográfica com ruído independente por observação se mostraram vulneráveis à reconstrução temporal, pois a garantia permanece válida por observação, mas o ataque explora a ausência de controle entre observações do mesmo ponto físico.

Para sistemas de vigilância municipal, recomendamos microagregação com $k{=}10$ como técnica primária. Quando uma garantia formal é estritamente requerida, por exigência contratual ou regulatória, DP com $\varepsilon \in [0{,}5; 1{,}0]$ surge como alternativa viável, desde que se esteja ciente da vulnerabilidade temporal sob agregação repetida. A publicação direta de coordenadas brutas, prática ainda adotada em diversos municípios, representa um risco desnecessário de reidentificação de imóveis privados, eliminável por técnicas simples e, neste estudo, sem perda detectável de utilidade.

As limitações encontradas foram, primeiro o uso de um único município e um único agente patogênico, onde os aspectos metodológicos são portáveis, mas os resultados quantitativos requerem replicação em outros contextos. Há também o escopo da tarefa preditiva, os resultados de utilidade referem-se a uma predição agregada em escala municipal, e a robustez observada decorre em parte dessa granularidade, conforme a Seção 7.4, não devendo ser extrapolada para tarefas de precisão fina, como a identificação de foco em raio de dezenas de metros, nas quais o deslocamento da anonimização é da mesma ordem da unidade de análise. O adversário avaliado é restrito à proximidade geométrica, os quatro ataques exploram distância *haversine* sem conhecimento semântico do território, onde bases públicas de pontos de interesse permitiriam restringir candidatos a locais compatíveis com o registro, e nesse caso generalização e microagregação podem se mostrar menos protetoras do que sugerem A1 e A4.

Quanto às escolhas de implementação, adotamos particionamento aleatório uniforme na microagregação, enquanto a literatura clássica [Domingo-Ferrer and Mateo-Sanz 2002] privilegia MDAV, que minimiza perda de informação. A cobertura temporal de 2023 é parcial, o CSV público termina em junho/2023 e os últimos *folds* amostram período de transição epidêmica, possivelmente não-estacionário, portanto as métricas são conservadoras. A generalização agressiva inviabiliza o *pipeline*, $d{=}1$ ($\sim$ 11 km) colapsa Porto Alegre em oito pontos, insuficiente para 15 micro-regiões. Já a microagregação não oferece garantia formal, a recomendação baseia-se em robustez empírica sob o modelo de adversário avaliado, e sob adversário caixa-branca ou múltiplas versões anonimizadas do mesmo dado pode degradar de forma que a privacidade diferencial não degradaria.

Como trabalhos futuros, identificamos quatro direções. A primeira é replicar o *framework* em outras capitais brasileiras, para isolar as propriedades estruturais do *pipeline* e verificar se a dominância de $k{=}10$ se mantém sob densidades urbanas distintas; a extensão da série de Porto Alegre para 2024 em diante permite ainda avaliar a estabilidade temporal dos resultados no mesmo município. Estender o modelo de adversário com ataques baseados em conhecimento semântico de pontos de interesse, como hospitais, escolas e afins, que reduzem o conjunto de candidatos plausíveis para além do que a distância *haversine* captura [Primault et al. 2019, Shokri et al. 2012]; investigar a privacidade diferencial local aplicada diretamente sobre contagens regionais agregadas, com potencial de garantia formal sem o custo temporal do mecanismo individual; e comparar a microagregação aleatória com o algoritmo MDAV [Domingo-Ferrer and Mateo-Sanz 2002], que diminui perda de informação por construção.

Agradecimentos

Os autores empregaram ferramentas de inteligência artificial para apoio à tradução de texto, revisão textual e formatação das tabelas e equações em código LaTeX. Este estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e grant #88887.954253/2024-00 (INCT ICoNIoT). Também foi financiado em parte pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (CNPq) - bolsas #314506/2023-3, #405940/2022-0 (INCT ICoNIoT) e #444978/2024-0. Também foi financiado em parte pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - bolsas #2020/05183-0 e #2023/00816-2.

Referências

Aleixo, R., Kon, F., Rocha, R., Camargo, M. S., and De Camargo, R. Y. (2022). Predicting dengue outbreaks with explainable machine learning. In *2022 22nd IEEE Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pages 940–947. IEEE.

Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., and Palamidessi, C. (2013). Geo-indistinguishability: Differential privacy for location-based systems. In *ACM SIGSAC Conference on Computer & Communications Security*, pages 901–914. ACM.

Botcrawl Security Reports (2025). Conasems data breach exposes 68,000 users and cpf numbers in major brazilian health sector leak. `https://botcrawl.com/conasems-data-breach/`. Acessado em maio de 2026.

Coelho, K., Okuyama, M., Nogueira, M., Vieira, A., Silva, E., and Nacif, J. (2024). Uma abordagem dinâmica para anonimização de dados de saúde por separatrizes. In

Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, pages 826–839, Porto Alegre, RS, Brasil. SBC.

Coelho, K., Okuyama, M., Nogueira, M., Vieira, A., Silva, E., and Nacif, J. (2025). Metodologia para avaliação da anonimização baseada em k-anonimato nos modelos de aprendizado de máquina. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 742–755, Porto Alegre, RS, Brasil. SBC.

de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., and Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3(1):1376.

Domingo-Ferrer, J. and Mateo-Sanz, J. M. (2002). Practical data-oriented microaggregation for statistical disclosure control. *IEEE Transactions on Knowledge and Data Engineering*, 14(1):189–201.

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

El Emam, K., Dankar, F. K., Issa, R., Jonker, E., Amyot, D., Cogo, E., Corriveau, J.-P., Walker, M., Chowdhury, S., and Vaillancourt, R. (2011). A systematic review of re-identification attacks on health data. *PLOS ONE*, 6(12):e28071.

Fraccaro, P., Beukenhorst, A., Sperrin, M., Harper, S., Palmier-Claus, J., Lewis, S., Van der Veer, S. N., and Peek, N. (2019). Digital biomarkers from geolocation data in bipolar disorder and schizophrenia: a systematic review. *Journal of the American Medical Informatics Association*, 26(11):1412–1420.

GDPR Register (2020). Brazil's health ministry's data leak. `https://www.gdprregister.eu/articles/brazils-health-ministrrys-data-leak/`. Acessado em maio de 2026.

Hoh, B., Gruteser, M., Xiong, H., and Alrabady, A. (2007). Preserving privacy in GPS traces via uncertainty-aware path cloaking. In *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS)*, pages 161–171. ACM.

Jin, F., Hua, W., Francia, M., Chao, P., Orlowska, M. E., and Zhou, X. (2023). A survey and experimental study on privacy-preserving trajectory data publishing. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):5577–5596.

Khanfor, A., Ghazzai, H., Yang, Y., and Massoud, Y. (2019). Application of community detection algorithms on social internet-of-things networks. In *2019 31st international conference on microelectronics (ICM)*, pages 94–97. IEEE.

Krumm, J. (2009). A survey of computational location privacy. *Personal and Ubiquitous Computing*, 13(6):391–399.

Lin, Y., Shen, Z., and Teng, X. (2021). Review on data sharing in smart city planning based on mobile phone signaling big data. *International Review for Spatial Planning and Sustainable Development*, 9(2):76–93.

Löbner, S., Tronnier, F., Pape, S., and Rannenberg, K. (2021). Comparison of de-identification techniques for privacy preserving data analysis in vehicular data sharing. In *ACM Computer Science in Cars Symposium*. ACM.

Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. In *22nd International Conference on Data Engineering (ICDE)*, pages 24–35. IEEE.

Majeed, A. and Lee, S. (2021). Anonymization techniques for privacy preserving data publishing: A comprehensive survey. *IEEE Access*, 9:8512–8545.

Murthy, S., Abu Bakar, A., Abdul Rahim, F., and Ramli, R. (2019). A comparative study of data anonymization techniques. In *Intl Conference on Big Data Security on Cloud (BigDataSecurity)*, pages 306–309.

Narayanan, A. and Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy*, pages 111–125. IEEE.

Park, S., Kim, R., Yoon, H., and Lee, K. (2021). Data privacy in wearable IoT devices: Anonymization and deanonymization. *Security and Communication Networks*, 2021(1):4973404.

Pimenta, I., Araújo, R., Rodrigues, R., Silveira, M., and Gomes, R. (2025). Anonimização de dados para inteligência artificial usando o algoritmo da tropa dos gorilas. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 448–461, Porto Alegre, RS, Brasil. SBC.

Primault, V., Boutet, A., Ben Mokhtar, S., and Brunie, L. (2019). The long road to computational location privacy: A survey. *IEEE Communications Surveys & Tutorials*, 21(3):2772–2793.

Ribeiro, S. E., Menezes, R. A., Portela, A. L., Araújo, T. P., and Gomes, R. L. (2023). Aplicando redes neurais e análise temporal para predição adaptativa de desempenho de rede. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, pages 490–503. SBC.

Saito, T. and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS ONE*, 10(3):e0118432.

Sampaio, S., Sousa, P. R., Martins, C., Ferreira, A., Antunes, L., and Cruz-Correia, R. (2023). Collecting, processing and secondary using personal and (pseudo)anonymized data in smart cities. *Applied Sciences*, 13(6).

Shokri, R., Theodorakopoulos, G., Le Boudec, J.-Y., and Hubaux, J.-P. (2012). Protecting location privacy: optimal strategy against localization attacks. *Proceedings of the ACM CCS*, pages 617–627.

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570.

Turgay, S., İlter, İ., et al. (2023). Perturbation methods for protecting data privacy: A review of techniques and applications. *Automation and Machine Learning*, 4(2):31–41.

Zang, H. and Bolot, J. (2011). Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 145–156. ACM.