

Análise Comparativa de Algoritmos de Alocação Escolar em Larga Escala

Anderson Pinheiro Cavalcanti^{1,2}, Samara Monteiro Xavier¹, Luiz Rodrigues³,
Péricles Miranda¹, Rafael Ferreira Mello^{1,2}

¹Universidade Federal Rural de Pernambuco (UFRPE)

²CESAR School

³Universidade Tecnológica Federal do Paraná (UTFPR)

{anderson.pinheiro, pericles.miranda}@ufrpe.br

{samara.xavier, rafael.mello}@ufrpe.br

luizrodrigues@utfpr.edu.br

Abstract. *Allocating students to public schools requires balancing preferences, capacity, and legal priority rules in a context of high demand and territorial inequality. This article compares Gale–Shapley, Multiple Queues, and an Adapted Constructive Heuristic through simulations with georeferenced synthetic data (zones A/B/C), up to 70,000 students, and three preferences per student. We assess runtime, preference fulfillment, the distribution by choice rank, and vulnerable students’ access to high-quality schools (Zone A) under equal weights and a pro-equity policy. Results show higher first-choice assignment under Multiple Queues, while Gale–Shapley preserves stability, and the pro-equity policy increases access to Zone A with diminishing gains in saturated scenarios.*

Resumo. *Alocar estudantes em escolas públicas exige conciliar preferências, capacidade e prioridades legais em um contexto de alta demanda e desigualdades territoriais. Este artigo compara Gale–Shapley, Filas Múltiplas e uma Heurística Construtiva Adaptada por meio de simulações com dados sintéticos georreferenciados (zonas A/B/C), até 70.000 estudantes e três preferências por aluno. Avaliam-se o tempo de execução, o atendimento às preferências, a distribuição por posição de escolha e o acesso de vulneráveis a escolas de alta qualidade (Zona A), com pesos iguais e política pró-equidade. Os resultados mostram maior alocação na 1ª opção com Filas Múltiplas, enquanto Gale–Shapley preserva a estabilidade e a pró-equidade amplia o acesso à Zona A com ganhos decrescentes em cenários saturados.*

1. Introdução

A alocação de estudantes nas escolas da rede pública de ensino é um dos desafios mais complexos e de maior impacto na gestão educacional contemporânea. Conhecido como Problema de Alocação de Alunos (PAA), ele transcende a simples logística de preenchimento de vagas, situando-se na interseção de políticas públicas de equidade, planejamento urbano e otimização de recursos [Bouzarth et al. 2018]. Anualmente, gestores públicos em âmbitos municipais e estaduais enfrentam a tarefa de distribuir centenas de milhares de alunos, buscando equilibrar múltiplos objetivos muitas vezes conflitantes, como a capacidade física das escolas, as preferências de alunos e suas famílias, os custos de

transporte e, fundamentalmente, os critérios de prioridade estabelecidos por lei, a fim de garantir o acesso justo à educação [Vianna et al. 2023].

O PAA é, fundamentalmente, um problema de otimização combinatória multiobjetivo pertencente à classe de complexidade NP-difícil [Vianna et al. 2023]. Por um lado, há a necessidade de respeitar as preferências das famílias, que buscam escolas com melhor reputação acadêmica, infraestrutura adequada ou maior segurança. Por outro lado, existem restrições operacionais severas, como o limite de alunos por sala de aula e a necessidade de minimizar os custos com transporte escolar, o que favorece a alocação baseada na proximidade geográfica (georreferenciamento). Adicionalmente, legislações nacionais, como o Estatuto da Criança e do Adolescente (ECA) no Brasil, impõem critérios de prioridade que devem ser estritamente observados, tais como a garantia de vaga para estudantes com deficiência e a preferência a irmãos na mesma unidade escolar.

A literatura acadêmica tem abordado o PAA por meio de diversas lentes metodológicas, o que reflete a interdisciplinaridade do problema. Uma vertente consolidada, oriunda da Pesquisa Operacional, modela o PAA como variante do Problema de Localização de Instalações, especificamente o problema das p -medianas, visando minimizar a distância total percorrida pela população estudantil [Alifi et al. 2017]. Outra vertente, fundamentada na Teoria dos Jogos e no Desenho de Mecanismos (*Market Design*), prioriza propriedades econômicas como a estabilidade do pareamento e a resistência à manipulação estratégica. Nesta categoria, o algoritmo de Aceitação Diferida (*Deferred Acceptance*) de Gale e Shapley (GS) tornou-se uma referência teórica por oferecer propriedades formais, como estabilidade e resistência à manipulação, garantindo que nenhum estudante perca uma vaga para outro com menor prioridade em uma escola que ambos preferem [Neto et al. 2022].

Apesar do avanço dessas técnicas e da riqueza de abordagens, há poucos estudos comparativos que analisem, sob as mesmas condições de dados e restrições, o desempenho de diferentes classes de algoritmos no PAA [Abdulkadiroğlu and Sönmez 2003, Pathak and Sönmez 2013, Vianna et al. 2023]. Essa lacuna é relevante, pois a escolha do método de alocação impacta diretamente indicadores de qualidade e justiça do processo, como a taxa de alocação na primeira opção, a distribuição equilibrada de vagas e a transparência dos critérios de prioridade [Hafalir et al. 2013, Abdulkadiroğlu and Sönmez 2003]. Diante desse cenário, este trabalho tem como objetivo principal realizar uma análise comparativa de três algoritmos de alocação de estudantes sob diferentes políticas de priorização:

- Gale–Shapley (*one-to-many* adaptado por *score*), representando métodos baseados em preferências e estabilidade;
- Filas Múltiplas, refletindo mecanismos administrativos com ordenação por prioridade [Abdulkadiroğlu et al. 2005];
- Heurística Construtiva Adaptada (inspirada em [Vianna et al. 2023]), que combina prioridade e satisfação em um processo de alocação escalável.

Ao comparar mecanismos de alocação sob políticas distintas em cenários realistas e congestionados, este trabalho fornece evidências quantitativas e um desenho experimental reproduzível para analisar *trade-offs* entre satisfação, aderência a prioridades e acesso de grupos vulneráveis a escolas de maior qualidade. Dessa forma, os resultados podem apoiar tanto o aprimoramento de plataformas de matrícula quanto a avaliação de regras

de negócio e políticas públicas, contribuindo para decisões mais informadas sobre como transformar critérios normativos em rotinas computacionais auditáveis e escaláveis. Assim, este artigo contribui para a comunidade de Informática na Educação com insumos que podem ter impacto direto nas políticas públicas educacionais ao aplicar métodos computacionais bem estabelecidos na análise do PAA.

2. Trabalhos Relacionados

A complexidade do PAA reside na necessidade de conciliar a eficiência no uso dos recursos públicos (vagas e transporte) com a equidade no acesso e com a satisfação das famílias. A literatura deste contexto pode ser segmentada em três grandes eixos: modelos de otimização exata focados em localização, abordagens de teoria dos jogos focadas em estabilidade, e métodos heurísticos focados em escalabilidade prática.

2.1. Otimização de Localização e Modelos Exatos

Uma abordagem clássica trata o PAA como uma extensão do problema de transporte ou de localização de instalações. O objetivo central é, frequentemente, a minimização da distância total percorrida pelos estudantes, assumindo que a proximidade é o principal proxy para o bem-estar e a eficiência de custos. [Alifi et al. 2017], por exemplo, aplicaram o modelo de p -medianas na cidade de Bandung, na Indonésia, utilizando Sistemas de Informação Geográfica (SIG) para otimizar o zoneamento escolar. O estudo reportou uma redução de 26,55% na distância total de deslocamento, evidenciando o potencial de economia de recursos. Similarmente, no Brasil, [Rodrigues and Mayerle 2014] propuseram um modelo de Programação Linear Inteira Mista (PLIM) para a rede de ensino de Manaus. O modelo inovou ao considerar não apenas o transporte, mas também os custos de abertura de novas turmas, tratando as escolas como instalações com capacidade elástica até certo limite.

Apesar da precisão matemática, modelos exatos enfrentam severas limitações de escalabilidade. [Mayerle et al. 2022] expandiram a formulação para incluir a alocação de professores e salas, mas reconheceram que a complexidade computacional cresce exponencialmente com o tamanho da rede, tornando a solução ótima inatingível em tempo hábil para redes estaduais de grande porte. Além disso, modelos puramente baseados em distância frequentemente falham em capturar nuances qualitativas, como a preferência das famílias por projetos pedagógicos específicos ou a necessidade de manter irmãos juntos [Cavalcanti et al. 2026].

2.2. Mecanismos de Mercado e Estabilidade

Em resposta às limitações dos modelos puramente logísticos, a literatura de *Market Design* destacou a importância das preferências e da estabilidade. O algoritmo de GS, ou *Deferred Acceptance*, é um dos mais importantes desta vertente. Sua propriedade fundamental é a estabilidade: não há par aluno-escola que prefira estar junto em detrimento de suas respectivas alocações atuais. [Fernandes 2018] analisou o sistema de matrículas de São Paulo e argumentou que a ausência de um mecanismo centralizado estável gerava ineficiências e percepção de injustiça. A simulação demonstrou que a adoção do GS poderia aumentar a transparência e reduzir a troca de escolas após a matrícula. No entanto, o GS não é isento de críticas. Em muitos cenários reais, ele pode resultar em uma taxa de alocação na primeira preferência inferior à de mecanismos de base (*naive*) como o de

Boston (ou *Immediate Acceptance*), que processa as escolhas de forma gulosa e imediata. O debate entre o mecanismo de Boston (maior satisfação imediata, mas manipulável) e o GS (estável e à prova de estratégia, mas potencialmente menos eficiente) é central na literatura recente [Cavalcanti et al. 2026].

[Arnosti 2023] e [Hermstrüwer 2025] exploram como a informação procedimental afeta a percepção de justiça desses algoritmos, sugerindo que, embora o GS seja teoricamente superior, o mecanismo de Boston pode ser percebido como mais justo por leigos devido à sua lógica de fila intuitiva. Ainda no contexto de estabilidade, [Neto et al. 2022] propuseram o algoritmo O2MSM para a alocação de locais de prova do ENEM. Trata-se de uma adaptação do casamento estável para um cenário *one-to-many*, focando na minimização de distâncias e maximização da ocupação, demonstrando a versatilidade do GS para além da matrícula escolar regular. Apesar disso, esses modelos são rígidos para grandes volumes de dados, o que limita sua aplicação prática em diversos cenários reais [Abdulkadiroğlu et al. 2005].

2.3. Heurísticas e Meta-heurísticas

Para lidar com a intratabilidade dos modelos exatos e a rigidez dos mecanismos de mercado em redes gigantescas, as heurísticas têm ganhado espaço. [Prima and Arymurthy 2018] e [Prima and Arymurthy 2019] aplicaram Algoritmos Genéticos e o Algoritmo Firefly para resolver o problema de localização-alocação em Jacarta, obtendo soluções subótimas de alta qualidade com baixo custo computacional. Estas abordagens permitem a incorporação de funções objetivo complexas e não lineares que seriam impossíveis em PLIM.

Também no cenário brasileiro, [Vianna et al. 2023] desenvolveram uma heurística construtiva específica para a Secretaria de Educação do Rio de Janeiro (SEEDUC-RJ). A solução prioriza a satisfação do aluno em grupos de prioridade, utilizando uma lógica gulosa ordenada. O estudo reportou uma melhoria de 9% a 15% na alocação em primeiras opções comparada ao sistema legado, com tempos de execução na ordem de minutos para centenas de milhares de alunos. Essa classe de algoritmos, que chamaremos de “Heurísticas Construtivas Adaptadas”, representa um meio-termo pragmático [Vianna et al. 2023]: busca-se a eficiência do mecanismo de Boston, mas com salvaguardas de prioridade que tentam mimetizar a justiça do GS, sem o custo computacional da estabilidade perfeita [Abdulkadiroğlu et al. 2006, Gale and Shapley 1962].

2.4. Lacuna e Questões de Pesquisa

A literatura indica que a maioria dos estudos concentra-se isoladamente em uma das dimensões: otimizam a logística (modelos exatos), garantem estabilidade (GS) ou buscam velocidade (heurísticas). Faltam estudos comparativos que coloquem essas abordagens “frente a frente” sob as mesmas restrições e dados, especialmente avaliando como elas reagem a políticas de equidade agressivas (pesos diferenciados para vulneráveis). Este artigo contribui ao comparar o desempenho do GS, de um mecanismo de Filas Múltiplas e da Heurística Construtiva em cenários controlados, analisando os *trade-offs* entre estabilidade, satisfação declarada e eficiência computacional. Para isso, foram definidas as seguintes Questões de Pesquisa (QP) que norteiam esse trabalho até seu objetivo principal:

QUESTÃO DE PESQUISA 1 (QP1):

Qual algoritmo maximiza a alocação nas primeiras opções e minimiza a exclusão?

QUESTÃO DE PESQUISA 2 (QP2):

Como diferentes algoritmos e pesos de prioridade impactam a entrada de alunos vulneráveis em escolas de alta qualidade?

3. Método

Esta seção descreve o desenho experimental utilizado para avaliar o desempenho comparativo dos algoritmos de alocação. A abordagem adotada é a de simulação computacional com dados sintéticos, permitindo o controle preciso das variáveis e a replicação de cenários [Law et al. 2007].

O PAA foi modelado como a atribuição de um conjunto de estudantes $S = \{s_1, s_2, \dots, s_n\}$ a um conjunto de escolas $E = \{e_1, e_2, \dots, e_m\}$, onde cada escola e_j possui uma capacidade c_j . Cada estudante possui uma lista ordenada de preferências P_s (escolas desejadas) e um vetor de atributos socioeconômicos que determinam sua prioridade. Com base em alguns trabalhos relacionados [Abdulkadiroğlu et al. 2005, Vianna et al. 2023, Arnosti 2023, Neto et al. 2022], foram implementados três algoritmos em Python para resolver este problema, detalhados a seguir.

Gale-Shapley Adaptado (GS): Esta é a implementação do algoritmo de Aceitação Diferida (*Student-Proposing Deferred Acceptance*). Os estudantes propõem a escola de sua maior preferência ainda não tentada. As escolas recebem as propostas e retêm provisoriamente os estudantes com os maiores *scores* de prioridade até o limite de sua capacidade, rejeitando os excedentes. Estudantes rejeitados tentam a próxima escola de sua lista na rodada seguinte. O processo repete-se até que não haja mais propostas ou vagas. O algoritmo garante estabilidade (ausência de par bloqueador) e é à prova de estratégia para os estudantes (revelar a verdadeira preferência é a melhor estratégia). O *score* de prioridade é calculado com base nos critérios da política vigente. Por exemplo, se um aluno possui 5 critérios de prioridade em uma política de pesos iguais, sua pontuação será 0,5.

Filas Múltiplas (FM): Simula processos administrativos baseados em rodadas de prioridade [Abdulkadiroğlu et al. 2005]. O algoritmo opera em rodadas sequenciais ($k = 1, 2, 3, \dots$). Na rodada k , são processadas apenas as k -ésimas opções de preferência de todos os alunos ainda não alocados. As escolas preenchem suas vagas disponíveis com os candidatos dessa rodada, ordenados por *score*, calculado da mesma forma que o algoritmo anterior. A alocação é definitiva (imediata): uma vez aceito, o aluno não é desalocado. O algoritmo tende a maximizar a alocação às primeiras preferências declaradas, pois atribui prioridade absoluta à ordem de escolha. No entanto, não é estável e penaliza severamente estudantes que não conseguem vaga na primeira opção, empurrando-os para escolas de última prioridade.

Heurística Construtiva Adaptada (HC): Com base na proposta de [Vianna et al. 2023], a HC é focada na eficiência e na satisfação. O algoritmo calcula um *score* global máximo para cada estudante (considerando seus atributos e a melhor combinação possível de critérios), com base nos multiplicadores de posição

propostos por [Vianna et al. 2023]. Por exemplo, se um aluno possui 3 critérios de prioridade e selecionou 3 escolas (A, B e C), sendo a escola A a primeira opção, ele terá um score máximo na escola A (multiplicador 25 x critérios 0,3 = 7,5 score final) em comparação com a escola B (multiplicador 20 x critérios 0,3 = 6,0 score final) e C (multiplicador 15 x critérios 0,3 = 4,5 score final) que terão valores menores. Ordena-se a lista global de estudantes de ordem decrescente por esse *score*. O algoritmo itera sobre essa lista, tentando alocar cada estudante à sua melhor opção de preferência disponível no momento em que é a sua vez. Se a primeira opção está cheia, tenta a segunda, e assim sucessivamente. Se nenhuma opção estiver disponível, o aluno vai para a lista de espera. A característica desse algoritmo é ser extremamente rápido ($O(N \log N)$ devido à ordenação) e de fácil implementação. Ele busca um equilíbrio entre respeitar a prioridade (score) e a preferência, sem as múltiplas iterações de propostas do GS nem a rigidez de rodadas do FM.

3.1. Conjunto de Dados e Cenários

Para garantir a validade externa do estudo, foi desenvolvido um gerador de dados (baseado em [Neto et al. 2022] que reproduz as características estruturais de redes públicas reais, incorporando correlações espaciais e socioeconômicas baseadas no Índice de Vulnerabilidade Social (IVS) [Costa et al. 2018]. O espaço geográfico simulado foi dividido em três zonas (A, B e C) que refletem o gradiente de vulnerabilidade urbana: a Zona C concentra a população de alta vulnerabilidade social ($\approx 60\%$ dos alunos vulneráveis), enquanto a Zona A concentra a população de menor vulnerabilidade ($\approx 60\%$ dos alunos não-vulneráveis). A Zona B é a zona intermediária que contempla aproximadamente 30% de vulneráveis e os mesmos 30% para não-vulneráveis. Além das zonas, definiu-se uma métrica para representar a qualidade da escola, com intervalos [0, 0,30] para a zona C, [0,31, 0,64] para a zona B e [0,65, 1,0] para a zona A. A população foi modelada em dois arquétipos principais: estudantes “vulneráveis” ($\approx 30\%$) e “não-vulneráveis” ($\approx 70\%$).

Foram definidos três cenários de escala crescente para avaliar o desempenho computacional e a eficácia das políticas (Tabela 1). Em todos os cenários, a oferta total de vagas foi configurada para ser ligeiramente inferior à demanda total (taxa de ocupação $\approx 1,00$), o que impulsionou a competição e permitiu a análise de alunos não alocados. Nos casos de escolas mais competitivas na Zona A, a demanda de primeira opção chegou a exceder a capacidade de vagas por mais de dez vezes, ou seja, se a escola possui 20 vagas, sua demanda chega a ser maior que 200 estudantes que colocaram essa escola como primeira opção.

Tabela 1. Métricas Descritivas dos Cenários de Teste

Cenário	Alunos	Escolas	Vagas	Taxa Ocup.	% Vuln.	Alu. A	Alu. B	Alu. C	Esc. A	Esc. B	Esc. C	Qual. μ
10k	10.000	100	9.959	1,004	29,52%	4.574	2.962	2.464	33	33	34	0,528
50k	50.000	500	49.785	1,004	29,94%	22.498	14.996	12.506	166	166	168	0,535
70k	70.000	700	70.032	1,000	30,15%	31.335	21.150	17.515	233	233	234	0,526

Nota: “Alu.” refere-se a alunos residentes na zona; “Esc.” às escolas localizadas na zona. “Qual. μ ” refere-se a Qualidade média

Foram simulados nove critérios de prioridade para a alocação. A definição dos critérios de prioridade é um componente central para a validade do modelo de simulação, pois determina como o algoritmo resolverá os conflitos de vagas nas unidades sobre-demanda. A estrutura adotada nesta pesquisa reflete o arcabouço jurídico brasileiro, integrando diretrizes do Estatuto da Criança e do Adolescente (ECA), do Plano Nacional de

Educação (PNE) e da legislação recente sobre proteção à mulher e à vulnerabilidade socioeconômica [COCK et al. 2022]. Os critérios de 1 a 7 consolidam dimensões de risco social e pessoal, permitindo que o sistema identifique e priorize crianças em situações de extrema fragilidade, como as que pertencem a famílias monoparentais ou cujas mães se encontram sob medidas protetivas. Os critérios 8 e 9, por sua vez, atendem a necessidades logísticas e de fortalecimento de vínculos familiares, garantindo o direito à proximidade residencial e à convivência de irmãos no mesmo ambiente escolar. Essa hierarquização garante que a alocação não seja meramente técnica, mas sim orientada por políticas públicas de equidade e proteção integral. A Tabela 2 detalha os critérios de prioridade.

Tabela 2. Definição dos Critérios de Prioridade para a Alocação Escolar

ID	Critério	Descrição e Aplicação no Modelo
1	Transferência de Renda	Priorização de famílias em situação de vulnerabilidade econômica, incluindo beneficiários de programas sociais.
2	Monoparentalidade	Reconhecimento da composição familiar monoparental como fator de vulnerabilidade social.
3	Criança com Deficiência	Prioridade absoluta para o atendimento de crianças com deficiência física, intelectual ou sensorial.
4	Violência Doméstica	Garantia de vaga próxima ao domicílio e sigilo para filhos de mulheres sob medida protetiva.
5	Medida de Proteção	Proteção integral para crianças que vivenciaram situações de violência intrafamiliar ou em acolhimento.
6	Medida Socioeducativa	Atendimento garantido aos filhos de adolescentes em cumprimento de medidas socioeducativas.
7	Mães sob Custódia	Acesso ao sistema de ensino visando o desenvolvimento integral de filhos de mulheres privadas de liberdade.
8	Irmãos na Escola	Direito de matrícula na mesma unidade escolar para irmãos que frequentam a mesma etapa básica.
9	Proximidade	Garantia de acesso à vaga em unidade escolar mais próxima da residência do estudante.

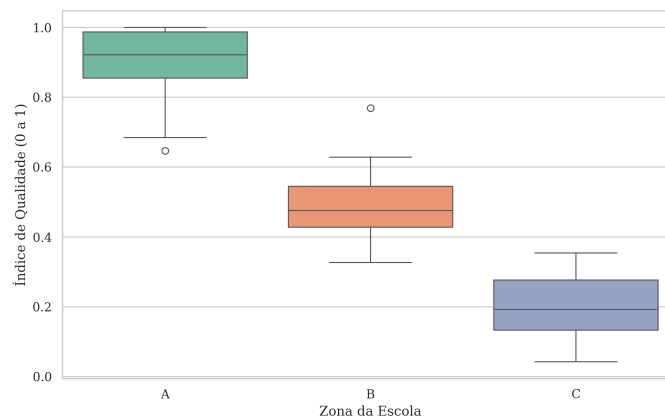
A Tabela 3 demonstra a distribuição da vulnerabilidade no conjunto de dados simulado. Enquanto a maioria da população (aproximadamente 63%) não atende a nenhum critério de prioridade social, existe um grupo minoritário, porém crítico, que acumula múltiplos fatores de risco. A maioria da população concentrou-se na faixa de 0 a 1 nos critérios de vulnerabilidade, com uma pequena parcela ($\approx 0,3\%$) apresentando alta prioridade acumulada (4 ou mais critérios).

As preferências dos alunos foram geradas por uma função de utilidade que pondera a qualidade da escola, a proximidade, a presença de irmãos e um fator aleatório idiossincrático. Cada estudante possui uma lista de três opções ordenadas da escola mais preferida à menos preferida. Esta modelagem gerou um cenário mais próximo da realidade: a qualidade média da primeira opção de escola para o grupo não-vulnerável situou-se em $\approx 0,95$, enquanto, para o grupo vulnerável, foi de $\approx 0,88$.

Como ilustrado na Figura 1, a infraestrutura escolar acompanha a segregação espacial. As escolas da Zona A apresentam índices de qualidade superiores e menor variância, tornando-se focos de extrema sobre-demanda. Em contraste, as zonas de maior vulnerabilidade apresentam maior dispersão na qualidade e índices médios inferiores.

Tabela 3. Distribuição da Quantidade de Critérios de Vulnerabilidade Acumulados por Cenário

Critérios Acumulados	Cenário 10k	Cenário 50k	Cenário 70k
0	6.301	31.251	43.779
1	2.554	12.801	17.841
2	919	4.663	6.624
3	198	1.141	1.550
4	24	141	200
5+	4	3	6
Total	10.000	50.000	70.000

**Figura 1. Distribuição da qualidade das escolas por zona geográfica**

3.2. Políticas de Priorização

Os algoritmos utilizam a seguinte função de *score* padronizado (S) para ordenar os candidatos na disputa por uma vaga:

$$S_i = \sum_{j=1}^9 (w_j \cdot x_{ij}) \quad (1)$$

onde $x_{ij} \in \{0, 1\}$ indica se o aluno i atende ao critério j , e w_j é o peso.

De acordo, foram utilizadas duas políticas nos experimentos. A Política A (Pesos Iguais) é uma abordagem afirmativa (Política B: Pró-Equidade), que privilegia critérios de vulnerabilidade social, que são definidos conforme o seguinte:

- **Política A (Pesos Iguais):** Todos os critérios têm peso $w = 1$. É o cenário de controle, onde proximidade e vulnerabilidade competem igualmente.
- **Política B (Pró-Equidade):** Critérios de vulnerabilidade (1-7) têm peso $w = 2$; os demais têm peso $w = 1$. Esta política simula uma ação afirmativa agressiva para garantir o acesso dos mais vulneráveis às escolas de sua escolha.

3.3. Métricas de Avaliação

Para avaliar os algoritmos e as políticas, a comparação considera métricas de satisfação de preferência, aderência às prioridades e desempenho computacional, com o objetivo

de oferecer subsídios para a escolha e o desenho de sistemas de alocação mais justos e eficientes no contexto da educação pública brasileira. Em particular, foram consideradas cinco métricas de avaliação: (1) Eficiência Computacional: tempo total de execução (em segundos); (2) Taxa de Alocação: percentual de alunos matriculados em alguma escola da lista; (3) Qualidade da Alocação (Satisfação): percentual de alunos alocados na 1ª, 2ª e 3ª preferências declaradas; (4) Satisfação Média: média ponderada da posição da preferência obtida (quanto menor, melhor, onde 1.0 seria o ideal); e (5) Porcentagem de alunos considerados vulneráveis alocados em escolas da zona A.

4. Resultados

Esta seção apresenta os resultados organizados em função das duas perguntas de pesquisa. Primeiramente, analisamos qual algoritmo/política maximiza a alocação nas primeiras opções e minimiza a exclusão dos estudantes dentro da lista declarada. Em seguida, investigamos como os algoritmos e os diferentes pesos de prioridade afetam o acesso de estudantes vulneráveis às escolas de alta qualidade da Zona A.

4.1. QP1: Qual algoritmo maximiza a alocação nas primeiras opções e minimiza a exclusão?

A Tabela 4 sintetiza o desempenho dos três algoritmos sob a Política A de pesos iguais. A leitura conjunta da alocação total, da posição média e da distribuição por preferência evidencia um *trade-off* entre a satisfação imediata e a redução da exclusão.

Tabela 4. Desempenho dos algoritmos na alocação total e na satisfação de preferências sob a Política A

Cenário	Algoritmo	Aloc. Total	Pos. Média	Exec. (s)	1ª Pref.	2ª Pref.	3ª Pref.
10k	Heurística	41,12%	1,65	0,67	56,08%	22,57%	21,35%
	Gale-Shapley	39,84%	1,58	0,69	61,47%	19,40%	19,13%
	Filas Múltiplas	39,18%	1,28	1,19	81,62%	8,32%	10,06%
50k	Heurística	26,54%	1,41	3,23	73,85%	11,61%	14,54%
	Gale-Shapley	26,22%	1,51	4,62	67,57%	14,02%	18,41%
	Filas Múltiplas	26,36%	1,18	5,93	87,53%	6,58%	5,89%
70k	Heurística	24,71%	1,36	4,67	77,17%	9,91%	12,93%
	Gale-Shapley	24,42%	1,43	6,17	72,82%	11,63%	15,55%
	Filas Múltiplas	24,58%	1,15	8,53	89,34%	5,97%	4,69%

Os resultados mostram que o algoritmo de Filas Múltiplas apresentou, em todos os cenários, a maior taxa de atendimento da 1ª preferência, alcançando 81,62%, 87,53% e 89,34% nos cenários de 10k, 50k e 70k, respectivamente. Esse comportamento também se reflete na melhor posição média de alocação, o que indica maior concentração de matrículas nas primeiras escolhas. Por outro lado, a Heurística apresentou a maior taxa de alocação total em todos os cenários, com 41,12%, 26,54% e 24,71%. Isso sugere maior capacidade de reduzir a exclusão na lista declarada, ainda que com menor concentração na 1ª opção do que o mecanismo de Filas Múltiplas. O Gale-Shapley ocupou uma posição intermediária. No cenário de 10k, superou a Heurística em 1ª preferência, mas nos cenários de 50k e 70k foi superado por ela. Além disso, apresentou maior dispersão entre a 2ª e a 3ª opção, o que indica um balanceamento menos concentrado na escolha principal. Em

síntese, os resultados mostram que não há um algoritmo dominante para todos os objetivos da RQ1. Se o foco estiver em maximizar a 1ª opção, Filas Múltiplas é o mecanismo mais eficaz. Se o objetivo for minimizar a exclusão e maximizar a alocação total na lista, a Heurística apresenta o melhor desempenho agregado.

Em termos de eficiência computacional, os resultados confirmam empiricamente o comportamento esperado a partir do desenho de cada algoritmo. A Heurística Construtiva apresentou os menores tempos de execução em todos os cenários, consistente com sua complexidade $O(N \log N)$ baseada em uma única ordenação global dos estudantes por score. O Gale-Shapley, por depender de múltiplas rodadas de propostas e rejeições até a convergência, apresentou tempos superiores e uma taxa de crescimento mais acentuada que a da Heurística à medida que o cenário se satura. Já as Filas Múltiplas registraram os maiores tempos em termos absolutos, o que sugere que o custo de recalculer e reordenar os candidatos por score a cada rodada supera o custo de uma ordenação global única realizada pela Heurística. Esses resultados reforçam que a Heurística Construtiva é a alternativa mais adequada para operações em larga escala sob restrições de tempo, sem comprometer de forma expressiva a qualidade da alocação.

4.2. QP2: Como diferentes algoritmos e pesos de prioridade impactam a entrada de alunos vulneráveis em escolas de alta qualidade (Zona A)?

A Tabela 5 apresenta o percentual de estudantes vulneráveis e não vulneráveis alocados em escolas da Zona A sob as políticas A e B.

Tabela 5. Acesso de estudantes vulneráveis e não-vulneráveis às escolas da Zona A

Cenário	Algoritmo	Política	Vuln. Zona A	Não-Vuln. Zona A
10k	Gale-Shapley	A (Igual)	70,36%	56,78%
		B (Equidade)	71,85%	55,08%
	Heurística	A (Igual)	68,93%	54,64%
		B (Equidade)	69,26%	54,15%
	Filas Múltiplas	A (Igual)	67,98%	61,13%
		B (Equidade)	68,53%	60,67%
50k	Gale-Shapley	A (Igual)	72,33%	64,79%
		B (Equidade)	73,23%	64,16%
	Heurística	A (Igual)	73,63%	62,77%
		B (Equidade)	73,55%	62,90%
	Filas Múltiplas	A (Igual)	70,62%	66,81%
		B (Equidade)	71,04%	66,25%
70k	Gale-Shapley	A (Igual)	70,33%	62,55%
		B (Equidade)	70,75%	61,93%
	Heurística	A (Igual)	70,75%	61,37%
		B (Equidade)	70,88%	61,12%
	Filas Múltiplas	A (Igual)	67,45%	65,32%
		B (Equidade)	67,80%	64,91%

Os resultados mostram que a Política B elevou o acesso de estudantes vulneráveis à Zona A na maior parte das combinações de cenário e algoritmo, embora com intensidade

limitada. No Gale-Shapley, os ganhos foram mais visíveis, especialmente no cenário de 10k, com aumento de 70,36% para 71,85%. Nos cenários de 50k e 70k, os incrementos foram menores, indicando uma redução do efeito da política à medida que a saturação aumenta. A Heurística apresentou, em parte dos cenários, os maiores níveis absolutos de acesso de vulneráveis à Zona A, como no cenário de 50k sob a Política A. No entanto, o ganho adicional com a Política B foi pequeno, sugerindo que parte do efeito distributivo já está incorporada à própria lógica de ordenação por *score*.

O algoritmo de Filas Múltiplas apresentou os menores níveis de acesso de vulneráveis à Zona A em todos os cenários. Mesmo com a Política B, os percentuais permaneceram inferiores aos dos demais mecanismos, o que indica menor sensibilidade redistributiva. Também se observa que o aumento do acesso de vulneráveis foi acompanhado de uma pequena redução no percentual de não-vulneráveis na Zona A, sinalizando que a política de pesos altera, ainda que moderadamente, a composição do grupo beneficiado pelas escolas de maior qualidade. Em conjunto, os resultados indicam que a entrada de vulneráveis na Zona A depende tanto da política adotada quanto da arquitetura do algoritmo. A Política B melhora a distribuição em quase todos os casos, mas seu efeito é mais pronunciado no Gale-Shapley e mais limitado nas Filas Múltiplas. A Heurística, por sua vez, combina bom desempenho distributivo com baixa perda de satisfação, o que a torna a alternativa mais equilibrada entre as dimensões analisadas.

5. Discussão

Esta seção interpreta os achados sob o ponto de vista de (i) por que a taxa de alocação em apenas três opções é estruturalmente baixa em cenários congestionados; (ii) os *trade-offs* entre satisfação imediata, estabilidade e incentivos estratégicos (QP1); e (iii) os limites de políticas pró-equidade baseadas em pesos frente à segregação espacial, em contraste com alternativas como reservas (QP2).

5.1. QP1: Satisfação de preferências versus estabilidade, prioridades e incentivos

Um achado dominante do estudo é a taxa de alocação relativamente baixa dentro das três preferências declaradas (24%–41%, dependendo do cenário). Esse resultado é coerente com a estrutura do problema simulada: (i) alta sobre-demanda concentrada nas escolas de maior qualidade (Zona A), e (ii) restrição de lista curta (apenas três opções), que limita a capacidade do mecanismo de encontrar uma alternativa viável antes de esgotar a lista. Em sistemas reais, a literatura mostra que restrições de lista e a falta de informação sobre as chances de admissão amplificam decisões defensivas e podem reduzir a taxa de alocação entre as escolhas informadas, tornando a métrica fortemente dependente do desenho institucional e do nível de congestionamento, e não apenas do algoritmo. Essa leitura é particularmente relevante quando a qualidade e a demanda são espacialmente correlacionadas, como sugerem os padrões de sobre-demanda observados no estudo, o que também aparece em trabalhos de planejamento e alocação que combinam qualidade, território e capacidade [Alifi et al. 2017, Rodrigues and Mayerle 2014, Mayerle et al. 2022, Vianna et al. 2023].

5.2. QP2: Análise da política Pró-equidade

Na QP2, dobrar pesos de vulnerabilidade (Política B) elevou de modo consistente, porém pequeno, o acesso de vulneráveis às escolas da Zona A (ganhos marginais decrescentes conforme a escala aumenta). Uma interpretação plausível é que a política de pesos

enfrenta a barreira espacial: a proximidade (Critério 9) tende a favorecer alunos residentes próximos às escolas de alta qualidade, enquanto alunos vulneráveis concentram-se em regiões distantes (Zona C) [Gortázar et al. 2023]. Assim, em unidades extremamente congestionadas, o incremento de peso pode ser insuficiente para superar a vantagem sistemática de proximidade, resultando em um ganho incremental limitado [Hafalir et al. 2013]. Esse padrão é consistente com a literatura de ação afirmativa em *school choice*. [Hafalir et al. 2013] mostra que políticas afirmativas baseadas em limites/ajustes suaves podem ser dominadas por reservas (*minority reserves*), nas quais parte das vagas é protegida para competição dentro do grupo prioritário antes da competição geral. Em cenários de segregação espacial e alta sobre-demanda, reservas tendem a produzir efeitos mais visíveis do que simples manipulações de pesos, justamente porque quebram a competição direta em condições estruturalmente desfavoráveis.

6. Limitações e Trabalhos Futuros

Apesar dos resultados apresentarem padrões consistentes entre cenários de escala (10k, 50k e 70k), este estudo possui limitações que devem ser consideradas na interpretação e na generalização dos achados. Os experimentos foram conduzidos com base em dados gerados por simulação a partir de uma única configuração do gerador (mesmas hipóteses estruturais de segregação espacial, distribuição de vulnerabilidade, qualidade das escolas e formação de preferências), variando-se essencialmente a escala do sistema. Assim, os achados devem ser interpretados como evidência em um cenário controlado, e não como estimativas diretas para uma rede específica. O conjunto de dados foi estruturado para conter apenas três opções de preferência por estudante. Essa escolha aumenta o realismo para sistemas que limitam o número de escolhas por simplicidade operacional, mas impõe um teto para a taxa de alocação “dentro da lista”, especialmente em ambientes congestionados. Portanto, parte da baixa taxa de alocação observada (aprox. 24%–41%) pode refletir a restrição de expressão de preferências, e não exclusivamente diferenças intrínsecas entre mecanismos.

Adicionalmente, o gerador de dados representa o espaço geográfico por meio de apenas três zonas (A, B e C), uma simplificação que captura o gradiente de qualidade e vulnerabilidade em termos agregados, mas que não reproduz a granularidade de uma rede real, na qual a proximidade entre aluno e escola é medida em distâncias contínuas e influenciada por barreiras de mobilidade urbana (malha viária, disponibilidade de transporte público, tempo de deslocamento), e não por pertencimento discreto a uma macrorregião.

Da mesma forma, o modelo não incorpora explicitamente os custos logísticos e sociais decorrentes da alocação de estudantes vulneráveis em escolas distantes de sua residência, tais como o tempo e o custo de deslocamento, a necessidade de transporte escolar subsidiado, ou o risco de evasão associado a trajetos mais longos ou inseguros.

Como trabalhos futuros, temos os seguintes direcionadores: (i) replicar os experimentos sob múltiplas parametrizações do gerador e, quando possível, em dados administrativos reais; (ii) variar sistematicamente o tamanho das listas de preferência; (iii) comparar políticas pró-equidade baseadas em pesos versus reservas; (iv) ampliar o conjunto de métricas com medidas explícitas de violação de prioridade e justiça procedimental; e (v) incorporar custos de deslocamento e de indicadores de risco de evasão como métricas complementares de avaliação.

Declaração sobre uso de Inteligência Artificial

Reconhecemos o uso de IA Generativa para auxiliar na elaboração e revisão deste artigo, que foi devidamente revisado, e assumimos total responsabilidade por seu conteúdo.

Referências

- Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2005). The boston public school match. *American Economic Review*, 95(2):368–371.
- Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2006). Changing the boston school choice mechanism: Strategy-proofness as equal access. Working Paper 11965, National Bureau of Economic Research.
- Abdulkadiroğlu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. *American economic review*, 93(3):729–747.
- Alifi, M. R., Hayati, H., and Supangkat, S. H. (2017). Optimization of school network using location-allocation analysis: Case study: Bandung, indonesia. In *2017 IEEE Region 10 Symposium (TENSYP)*, pages 1–6. IEEE.
- Arnosti, N. (2023). Lottery design for school choice. *Management Science*, 69(1):244–259.
- Bouzarth, E. L., Forrester, R., Hutson, K. R., and Reddoch, L. (2018). Assigning students to schools to minimize both transportation costs and socioeconomic variation between schools. *Socio-Economic Planning Sciences*, 64:1–8.
- Cavalcanti, A. P., Xavier, S. M., da Silva Venâncio, J. M., de Miranda, P. B. C., do Nascimento, A. C. A., and de Mello, R. F. L. (2026). Sistemas de alocação de estudantes em escolas: Uma revisão sistemática da literatura. *Revista Brasileira de Informática na Educação*, 34:458–477.
- COCK, J. C. A., NASCIMENTO, D., NASCIMENTO, A. M., COSTA, P. A., BONAMINO, A., and DE, M. C. (2022). Pesquisas sobre implementação de políticas educacionais no brasil: um estado do conhecimento. *Educação em Revista*, 38.
- Costa, M. A., dos Santos, M. P. G., Marguti, B., Pirani, N., Pinto, C. V. d. S., Curi, R. L. C., Ribeiro, C. C., and de Albuquerque, C. G. (2018). Vulnerabilidade social no brasil: conceitos, métodos e primeiros resultados para municípios e regiões metropolitanas brasileiras. Technical report, Texto para Discussão.
- Fernandes, G. A. d. A. L. (2018). O problema das matrículas escolares: uma discussão sobre a melhoria das regras utilizadas em são paulo. *Revista de Administração Pública*, 52(4):740–752.
- Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American mathematical monthly*, 69(1):9–15.
- Gortázar, L., Mayor, D., and Montalbán, J. (2023). Residence-based priorities and school choice. *Economics of Education Review*, 95:102384.
- Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2013). Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363.

- Hermstrüwer, Y. (2025). Procedural information as a “game changer” in school choice. *Games*, 16(3):25.
- Law, A. M., Kelton, W. D., and Kelton, W. D. (2007). *Simulation modeling and analysis*, volume 3. Mcgraw-hill New York.
- Mayerle, S. F., Rodrigues, H. F., de Figueiredo, J. N., and Chiroli, D. M. D. G. (2022). Optimal student/school/class/teacher/classroom matching to support efficient public school system resource allocation. *Socio-Economic Planning Sciences*, 83:101341.
- Neto, G. F., Miranda, P. B., Mello, R. F., and Nascimento, A. C. (2022). A novel one-to-many matching method for the assignment problem: An enem case study. In *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, pages 164–173. SBC.
- Pathak, P. A. and Sönmez, T. (2013). School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1):80–106.
- Prima, P. and Arymurthy, A. M. (2018). Optimization of school location-allocation using genetic algorithm. In *2018 8th International Workshop on Computer Science and Engineering, WCSE 2018*, pages 750–755. International Workshop on Computer Science and Engineering (WCSE).
- Prima, P. and Arymurthy, A. M. (2019). Optimization of school location-allocation using firefly algorithm. In *Journal of Physics: Conference Series*, volume 1235, page 012002. IOP Publishing.
- Rodrigues, H. F. and Mayerle, S. F. (2014). Modelo capacitado para alocação de alunos em redes públicas de ensino. *XLVI Simpósio Brasileiro de Pesquisa Operacional*.
- Vianna, C. R., Teles, C. A., Pontes, R. S., Belloze, K. T., Cirto, F. E., Brandão, D., de Assis, L. S., and de Carvalho, F. F. (2023). Heuristic for the students allocation in brazilian public schools: a case study. In *2023 18th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–7. IEEE.