

УДК 519.6:004.89

85. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЖАДНОГО И ГЕНЕТИЧЕСКОГО АЛГОРИТМОВ В ЗАДАЧЕ ОПТИМИЗАЦИИ РАСПРЕДЕЛЕНИЯ ВМЕСТИМОСТИ ДЛЯ ФИКСИРОВАННЫХ МАРШРУТОВ КАРПУЛИНГА

Али Марван А.Ф.А., магистрант гр.476701

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Федосенко В.А., кандидат технических наук, доцент

Аннотация. В статье проведен сравнительный анализ жадного алгоритма (Best-Fit) и генетического алгоритма (GA) в задаче распределения мест для междугороднего карпулинга. Установлено, что генетический алгоритм увеличивает число обслуженных пассажиров на 24%, что напрямую конвертируется в рост выручки оператора. Несмотря на высокую вычислительную сложность, GA позволяет реализовать полноценную модель управления доходами (Revenue Management). Авторами предложен переход от простой оптимизации к экономической модели, включающей динамическое ценообразование и гибкие скидки до 30% для минимизации потерь от незаполненных мест.

Ключевые слова. Карпулинг, генетический алгоритм, управление доходами, оптимизация выручки, динамическое ценообразование, многомерная задача о рюкзаке.

В современных системах совместной мобильности эффективность использования автопарка напрямую зависит от алгоритмов сопоставления водителей и пассажиров. Оптимальное распределение ресурсов в данном контексте выступает ключевым фактором экономической устойчивости сервиса, позволяя максимизировать выручку оператора за счет снижения доли пустых мест и одновременно минимизировать экологическую нагрузку на транспортную сеть через повышение удельной загрузки каждого транспортного средства. В то время как классический райдхейлинг опирается на динамическую маршрутизацию по запросу, междугородской карпулинг и кастомизированные шаттлы оперируют в условиях фиксированного маршрута и расписания. В таких условиях фокус смещается с задачи построения пути на комбинаторную задачу назначения подмножества пассажиров на ограниченное количество посадочных мест транспортного средства.

Вычислительная сложность поиска оптимального распределения мест в мультисегментных маршрутах относится к классу NP-трудных задач [3]. Использование точных методов (например, симплекс-метода или алгоритма ветвей и границ) для крупномасштабных систем признается нерациональным из-за «комбинаторного взрыва». В связи с этим в области транспортной оптимизации доминируют эвристические и метаэвристические подходы [2]. Поиск баланса между скоростью вычислений и качеством решения имеет критическое значение для внедрения систем динамического ценообразования и оперативного управления доходами платформы. Целью данной работы является теоретический и эмпирический сравнительный анализ детерминированной эвристики (жадного алгоритма Best-Fit) и стохастического глобального поиска (генетического алгоритма) в задаче статической аллокации пассажиров

Математическая постановка задачи. Рассматриваемая задача концептуально сводится к многомерной 0–1 задаче о рюкзаке (Multidimensional Knapsack Problem, MKP) [2].

Пусть задан фиксированный маршрут транспортного средства, состоящий из последовательности остановок $1, 2, \dots, m$. Транспортная система получает множество заявок на перевозку $J = \{1, \dots, n\}$. Каждая заявка $j \in J$ характеризуется станцией посадки o_j , станцией высадки d_j (где $o_j < d_j$) и запрашиваемым пассажиро-километражем t_j . Вводится бинарная переменная решения $x_j \in \{0, 1\}$, где $x_j = 1$, если заявка j принимается к обслуживанию, и $x_j = 0$ в случае отказа [2].

Ключевым ограничением модели является ограничение вместимости (1): на каждом сегменте маршрута суммарное число пассажиров в салоне не должно превышать физической вместимости транспортного средства:

$$\sum x_j \leq C, \forall k \in \{1, \dots, m-1\}, \text{ где } j: o_j \leq k < d_j \quad (1)$$

где x_j — бинарная переменная принятия заявки j (1 — принята, 0 — отклонена); C — максимальная вместимость транспортного средства (число мест); k — индекс сегмента маршрута между остановками k и $k+1$; o_j и d_j — остановки посадки и высадки пассажира j соответственно; m — общее число остановок на маршруте.

Целевая функция (2) максимизирует общее число обслуженных пассажиров: каждая принятая заявка вносит вклад 1 в целевое значение независимо от длины поездки:

$$\max \sum x_j, j \in \{1, \dots, n\} \quad (2)$$

где x_j — бинарная переменная решения для заявки j ; n — общее число поступивших заявок на перевозку; суммирование ведётся по всем $j \in J = \{1, \dots, n\}$.

В многокритериальных постановках данная функция может быть расширена для учёта пассажиро-километража ($\max \sum t_j x_j$).

Алгоритмические решения. Жадный алгоритм Best-Fit принимает локально оптимальные решения на каждом шаге без применения механизмов возврата [1]. В реализованном варианте пассажиры сортируются по убыванию длины поездки — от наиболее длинных к коротким. Данный критерий обоснован тем, что длинные поездки занимают наибольшее число сегментов и потребляют наибольший объём провозной ёмкости, определяемой ограничением (1).

Стратегия Best-Fit выбрана среди трёх реализованных вариантов (FirstFit, BestFit, RandomFit) ввиду её наилучших теоретических характеристик: в отличие от FirstFit, она минимизирует остаточную ёмкость на «узком» сегменте, назначая пассажира в транспортное средство с наименьшим свободным местом. Вычислительная сложность ограничена $O(N \log N)$, что гарантирует высокую масштабируемость. Однако метод имеет существенный разрыв оптимальности: приоритетный выбор длинных маршрутов может заблокировать ёмкость для комбинации коротких поездок [2].

Генетический алгоритм (GA) преодолевает ограничения локальных оптимумов посредством имитации популяционной эволюции. Применяется перестановочное кодирование: хромосома — вектор $\pi = (\pi_1, \pi_2, \dots, \pi_N)$ очередности обработки N заявок [3]. Качество каждой перестановки оценивается функцией приспособленности (3), вычисляемой детерминированным декодером, который последовательно применяет ограничение (1) к каждой заявке в порядке, заданном хромосомой.

Функция приспособленности реализована в виде (3) — взвешенной суммы двух критериев, обеспечивающей абсолютный приоритет числа обслуженных пассажиров над суммарным расстоянием:

$$F(\pi) = 10\,000 \cdot \sum x_{\pi i} + \sum t_i x_{\pi i} \quad (3)$$

где π — хромосома (перестановка), задающая порядок обработки заявок декодером; $x_{\pi i}$ — бинарная переменная, определяемая детерминированным декодером: равна 1, если заявка на позиции i в перестановке π успешно размещена в транспортном средстве, и 0 в противном случае; t_j — пассажиро-километраж заявки j (длина поездки в сегментах: $d_j - o_j$); коэффициент 10000 обеспечивает абсолютный приоритет числа обслуженных пассажиров над суммарным расстоянием.

Первый слагаемый с коэффициентом 10000 обеспечивает абсолютный приоритет максимизации числа обслуженных пассажиров. Второй слагаемый (суммарный пассажиро-километраж) выступает в роли тай-брейкера: при равном числе пассажиров предпочтение отдаётся решению с большим суммарным расстоянием. Для обеспечения генетического разнообразия и предотвращения деградации валидных перестановок применяется оператор частично отображаемого кроссовера (PMX) совместно с адаптивной мутацией трёх типов: обмен позиций (Swap), перемещение элемента (Relocate) и реверс подпоследовательности (Reverse). Начальная популяция включает две детерминированные затравки — сортировку по возрастанию и убыванию длины поездки — и случайные перестановки. Асимптотическая временная сложность GA масштабируется как $O(P \cdot G \cdot N)$, где P — размер популяции, G — количество поколений [3].

Вычислительный эксперимент и результаты. Эксперимент реализован в виде Python-скрипта сравнительного тестирования. Для каждого из четырёх сценариев формировалась воспроизводимая случайная задача с фиксированным начальным значением генератора ($\text{seed} = 42$). Вместимость водителей задавалась случайно в диапазоне 2–4 места; маршруты пассажиров генерировались равномерно по всему диапазону остановок. Жадный алгоритм запускался однократно; результаты ГА усреднялись по трём независимым запускам ($P = 80$, $G = 100$) для снижения влияния стохастичности. Нагрузка в каждом сценарии намеренно превышает провозную ёмкость на 90–110%: при такой нагрузке не все пассажиры могут быть обслужены, поэтому порядок их рассмотрения — и, следовательно, выбор алгоритма — непосредственно влияет на итоговое число принятых заявок. Заполняемость вычисляется по формуле: (сумма пассажиро-сегментов назначенных пассажиров) / (общее число сегментов всех транспортных средств) $\times 100\%$. Результаты представлены в Таблице 1.

Таблица 1. Сравнительный анализ эффективности алгоритмов распределения пассажиров.

Сценарий	Алгоритм	Обслужено пассажиров	Дистанция (км)	Заполняемость (%)	Время (мс)
Low_Demand	Жадный	32 / 45	69.0	92.0	0.1
Low_Demand	Генетический	40 / 45	69.0	92.0	804.0
Medium_Urban	Жадный	97 / 130	324.0	80.0	0.4
Medium_Urban	Генетический	119 / 130	324.3	80.1	4432.0
High_Congestion	Жадный	183 / 250	872.0	86.5	1.4
High_Congestion	Генетический	221 / 250	854.0	84.7	11511.0
Extreme_Peak	Жадный	310 / 450	1879.0	83.1	4.1
Extreme_Peak	Генетический	385 / 450	1739.3	76.9	33384.0

Число обслуженных пассажиров. Во всех четырёх сценариях GA стабильно превосходит жадный алгоритм по основному критерию — числу принятых заявок. В сценарии Low_Demand GA обслуживает 40 из 45 пассажиров против 32 у жадного (+25%). В Medium_Urban — 119 против 97 (+22.7%). В High_Congestion — 221 против 183 (+20.8%). Наиболее выраженный эффект наблюдается в Extreme_Peak: GA принимает 385 заявок против 310 у жадного алгоритма, что соответствует приросту в 75 пассажиров (+24.2%). Это объясняется тем, что при большом числе остановок и высокой нагрузке пространство допустимых перестановок значительно расширяется и эволюционный поиск эффективнее исследует его по сравнению с детерминированной жадной стратегией.

Парадокс заполняемости. Несмотря на превосходство по числу посадок, GA демонстрирует меньшую заполняемость салона в High_Congestion и Extreme_Peak. В Extreme_Peak жадный алгоритм достигает 83.1%, тогда как GA — лишь 76.9%, хотя и обслуживает на 75 пассажиров больше. Это явление обусловлено структурой целевой функции (3): метаэвристика обнаруживает, что для максимизации $\sum x_j$ выгоднее разместить нескольких пассажиров с короткими поездками там, где один длинный маршрут занимал бы то же место. Каждое место используется «дробно» — несколькими короткими поездками вместо одной длинной, — что снижает суммарный пассажиро-километраж и, следовательно, метрику заполняемости. Жадный алгоритм, сортирующий по убыванию длины поездки, напротив, стремится закрыть сквозные маршруты и генерирует больший пассажиро-километраж.

Время выполнения. Разрыв во времени выполнения носит принципиальный характер. Жадный алгоритм обрабатывает 45 заявок за 0.1 мс, 450 — за 4.1 мс, что соответствует теоретической сложности $O(N \log N)$ и делает его пригодным для систем реального времени. GA требует 804 мс (Low_Demand), 4432 мс (Medium_Urban), 11511 мс (High_Congestion) и 33 384 мс (Extreme_Peak) — в сотни и тысячи раз больше. Это исключает применение GA в системах с жёсткими временными ограничениями, однако делает его эффективным инструментом стратегического планирования: при формировании расписания межгородских маршрутов в ночное время задержка в десятки секунд приемлема в обмен на существенное увеличение числа обслуженных пассажиров.

Масштабируемость преимущества GA. С ростом размера задачи абсолютный прирост GA увеличивается: +8 пассажиров (Low_Demand), +22 (Medium_Urban), +38 (High_Congestion), +75 (Extreme_Peak). Это свидетельствует о том, что при большей размерности задачи пространство поиска становится богаче и эволюционный подход полнее реализует свой потенциал, тогда как жадный алгоритм демонстрирует стабильную, но ограниченную эффективность, не способную адаптироваться к глобальной структуре задачи.

Заключение. Проведённый анализ показывает, что генетический алгоритм обслуживает на 24% больше пассажиров, чем жадный алгоритм Best-Fit, за счёт глобального поиска в пространстве перестановок, тогда как жадный алгоритм выигрывает по скорости (4.1 мс против 33 секунд) и заполняемости салона. С экономической точки зрения выбор алгоритма непосредственно влияет на выручку оператора: каждое дополнительно занятое место генерирует доход, а незаполненный салон — прямые потери. Данная работа представляет собой первый этап исследования; следующим шагом является интеграция модуля динамического ценообразования, в котором базовый тариф определяется регрессионной моделью по факторам расстояния, спроса и типа дня, а затем корректируется в зависимости от текущей заполняемости — вплоть до 30% скидки при пустом салоне. Такой подход

позволит оценить не только число обслуженных пассажиров, но и совокупную выручку оператора, превращая задачу комбинаторной оптимизации в полноценную экономическую модель управления доходами карпулинговой платформы.

Список использованных источников:

1. *Introduction to Algorithms* / T. H. Cormen [et al.] // 4th ed. — Cambridge : MIT Press, 2022.
2. *Heuristic and metaheuristic algorithms for the multidimensional knapsack problem: A comprehensive review* / Y. Wang, X. Li // *Computers & Operations Research*, 2023. — Vol. 152. — Art. 106132. — DOI: 10.1016/j.cor.2023.106132.
3. *Solving Multidimensional Knapsack Problem with Bayesian Multiploid Genetic Algorithm* / E. Gazioglu // *Journal of Soft Computing and Artificial Intelligence*, 2022. — Vol. 3, No. 2. — P. 58–64.

UDC 519.6:004.89

COMPARATIVE ANALYSIS OF GREEDY AND GENETIC ALGORITHMS IN THE CAPACITY ALLOCATION OPTIMIZATION PROBLEM FOR FIXED CARPOOLING ROUTES.

Ali Marwan A.F.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Fedosenko V.A. – PhD in Technical Sciences, Associate Professor

Annotation. This paper evaluates a greedy algorithm (Best-Fit) and a genetic algorithm (GA) for seat allocation in intercity carpooling. The results show that the GA increases the number of passengers served by 24%, directly impacting the operator's profitability by reducing the "lost opportunity" of empty seats. While the greedy algorithm ($O(N \log N)$) is faster, the GA enables a sophisticated Revenue Management approach. The study concludes that integrating such algorithms with dynamic pricing—offering up to 30% discounts based on occupancy—transforms combinatorial optimization into a robust economic model for shared mobility platforms.

Keywords. Carpooling, genetic algorithm, revenue management, profit optimization, dynamic pricing, multidimensional knapsack problem.