

УДК 004.89

Худяков Александр Кириллович, магистрант, МИРЭА – Российский технологический университет, Москва

**МЕТОДЫ ОПТИМИЗАЦИИ И МАШИННОГО ОБУЧЕНИЯ В
ЗАДАЧАХ ДИНАМИЧЕСКОЙ МАРШРУТИЗАЦИИ ГРУЗОВЫХ
ПОТОКОВ: ОБЗОР СОВРЕМЕННОГО СОСТОЯНИЯ**

Аннотация

В статье рассмотрены методы и инструменты динамической маршрутизации грузовых потоков. Три класса подходов: точные методы оптимизации (MILP, алгоритм минимального стоимостного потока), эвристические методы (жадные алгоритмы, генетические алгоритмы, метаэвристики) и методы машинного обучения (градиентный бустинг, рекуррентные нейронные сети, трансформеры). Для каждого класса разобраны вычислительная сложность, масштабируемость и пригодность для задач магистральной логистики. Дан сравнительный анализ инструментов: OR-Tools, IBM CPLEX, LightGBM, XGBoost, Prophet, LSTM. Отдельно рассмотрена интеграция прогнозирования и оптимизации в единый конвейер с примером реализации на реальных логистических данных.

Annotation

This paper reviews methods and tools for dynamic freight routing. Three classes are covered: exact optimization methods (mixed-integer linear programming, min-cost flow), heuristic methods (greedy algorithms, genetic algorithms, metaheuristics), and machine learning methods (gradient boosting, recurrent neural networks,

transformers). For each class, the paper examines computational complexity, scalability, and fit for trunk logistics. Software tools are compared: OR-Tools, IBM CPLEX, LightGBM, XGBoost, Prophet, LSTM. The paper also covers the integration of forecasting and optimization into a single pipeline, with an implementation example on real logistics data.

Ключевые слова: маршрутизация грузовых потоков, задача о транспортировке, Min-Cost Flow, MILP, LightGBM, градиентный бустинг, OR-Tools, прогнозирование спроса, динамическое планирование, электронная коммерция.

Keywords: freight flow routing, transportation problem, Min-Cost Flow, MILP, LightGBM, gradient boosting, OR-Tools, demand forecasting, dynamic planning, e-commerce.

Введение

Маршрутизация грузовых потоков между складскими узлами – одна из ключевых операционных задач в электронной коммерции. Сложность определяется не только числом узлов и направлений, но и нестабильностью спроса: сезонные пики, акции и нерегулярные события способны за короткое время в два-три раза увеличить объём потоков на отдельных направлениях.

Большинство действующих промышленных систем строят маршрутные схемы по усреднённым историческим данным и пересчитывают их раз в несколько дней или недель. При переменном спросе это даёт предсказуемый результат: одни направления перегружены, другие недозагружены.

За последнее десятилетие в задачах маршрутизации сложились три направления: точные методы оптимизации с гарантией глобального оптимума, эвристические методы с приемлемым временем работы и методы машинного обучения, адаптирующие схему к текущему состоянию спроса. На практике лучшие результаты получают при совместном применении: прогностическая модель готовит входные данные для оптимизатора.

Цель статьи – систематизировать методы и инструменты динамической маршрутизации грузовых потоков, разобрать их сильные стороны и

ограничения, а также рассмотреть подходы к совместной работе прогнозирования и оптимизации.

Классические методы оптимизации маршрутов

Классификация методов планирования маршрутов приведена на Рисунке 1. Исторически первым изученным классом стали точные методы с гарантией глобального оптимума. Среди них три основных подхода, различающихся постановкой задачи и вычислительной сложностью.



Рисунок 1 – Классификация методов планирования маршрутов грузовых потоков

Задача маршрутизации транспортных средств (Vehicle Routing Problem, VRP) – один из наиболее изученных классов комбинаторной оптимизации [1]. В базовой постановке требуется найти маршруты для парка транспортных средств, обслуживающих клиентов с минимальными суммарными затратами. Расширения VRP включают ограничения вместимости (CVRP), временные окна (VRPTW), многодепотную структуру (MDVRP). Задача NP-hard: точное решение на крупных экземплярах требует экспоненциального времени при наивном переборе. На практике используют коммерческие MILP-решатели – IBM CPLEX, Gurobi, – сочетающие метод ветвей и границ с эвристическими стратегиями отсечения. На задачах с сотнями узлов время работы таких решателей исчисляется часами, а требования к оперативной памяти превышают 100 ГБ [2].

Задача смешанного целочисленного линейного программирования (MILP) – обобщение VRP, допускающее произвольные линейные ограничения с целочисленными переменными. Это наиболее гибкий инструмент моделирования, однако его вычислительная стоимость делает оперативный пересчёт схемы практически нереализуемым.

MCF – это частный случай линейного программирования на сети. В отличие от MILP, он точно решается за полиномиальное время. Формулировка задачи стандартная: каждому ребру (i, j) назначены стоимость единицы потока $c(i, j)$ и пропускная способность $u(i, j)$, каждому узлу – баланс $b(i)$, положительный для источников и отрицательный для стоков. Нужно найти такое распределение потоков $x(i, j)$, которое минимизирует суммарную стоимость при соблюдении всех балансов и пропускных способностей. Алгоритм последовательного кратчайшего пути (SSP) справляется с этим за $O(n \cdot |E| \cdot \log|V|)$, где n – суммарный поток [3]. На сети из нескольких сотен узлов это занимает меньше секунды на обычном компьютере. Готовая реализация MCF есть в OR-Tools от Google.

MCF хорошо подходит для магистральной логистики, где задача – распределить консолидированные грузовые потоки по межскладской сети, не привязываясь к конкретным машинам и расписаниям. Это и есть принципиальная разница с VRP: VRP работает на оперативном уровне и решает, какой заказ едет на какой машине.

Эвристические методы

Эвристические методы нужны там, где точное решение просто не успевает уложиться в допустимое время. Глобального оптимума они не обещают, зато дают достаточно приличный результат за секунды или минуты. Жадные алгоритмы строят решение пошагово, выбирая на каждом шаге локально оптимальный вариант. В маршрутизации типичный пример – последовательное добавление ближайшего непосещённого узла. Метод прост и

быстр, но качество сильно зависит от структуры задачи и нередко уступает оптимуму на 20–40% [1].

Генетические алгоритмы имитируют биологическую эволюцию: начальная популяция допустимых решений обрабатывается операторами скрещивания и мутации, отбор оставляет решения с лучшим значением целевой функции. Маршрут традиционно кодируется как перестановка узлов. Метод хорошо параллелизуется и устойчив к локальным минимумам, но требует тщательной настройки гиперпараметров и нескольких тысяч итераций для сходимости [4].

Метаэвристики – муравьиные колонии (АСО) и пчелиные алгоритмы – имитируют коллективное поведение биологических агентов. В АСО каждый «муравей» строит маршрут по феромонному следу, усиливающемуся на коротких путях и испаряющемуся со временем. АСО успешно применяется к различным вариантам VRP, по качеству сопоставимо с генетическими алгоритмами при меньшем числе параметров [5].

Общий недостаток эвристик применительно к магистральной логистике – отсутствие гарантии качества и сложность динамического обновления прогнозных данных о спросе в качестве входных параметров.

Методы машинного обучения для прогнозирования спроса

Прогнозирование спроса на перевозки напрямую определяет качество входных данных для оптимизатора, а значит – и качество итоговой маршрутной схемы. Рассмотрим основные классы методов.

Статистические методы – ARIMA и её расширения (SARIMA, ARIMAX) – долго оставались стандартом для прогнозирования временных рядов спроса. Достоинства: интерпретируемость, строгая математическая основа, устойчивость на небольших выборках. Недостаток применительно к логистике с тысячами направлений – необходимость строить отдельную модель для каждого ряда; при масштабе в десятки тысяч маршрутов это вычислительно неэффективно [6].

Prophet – раскладывает временной ряд на трендовую, сезонную и событийную компоненты. Хорошо улавливает праздничные эффекты и нерегулярные всплески, проста в настройке. Но, как и ARIMA, работает с одним рядом за раз – это ограничивает применимость при большом числе направлений [7].

Градиентный бустинг – в первую очередь LightGBM и XGBoost – прочно занял место стандарта для прогнозирования спроса на табличных данных. Ключевое практическое преимущество: одна модель покрывает сразу все направления. Пары складов кодируются как категориальные признаки, и статистика с загруженных маршрутов «перетекает» к слабонагруженным через общие признаки – как раз там, где истории мало. LightGBM дополнительно ускоряет обучение через алгоритм GOSS (Gradient-based One-Side Sampling): при построении каждого дерева объекты отбираются взвешенно, а не поровну [8]. По независимым бенчмаркам бустинг стабильно выигрывает у нейросетей на табличных данных при сопоставимом времени обучения.

LSTM и GRU теоретически способны улавливать долгосрочные зависимости в рядах – и это их главный аргумент. На практике в логистическом прогнозировании картина менее радужная: данных по каждому направлению нужно заметно больше, обучение занимает на порядок дольше, а разобраться в том, почему модель выдала именно такой прогноз, значительно сложнее [9]. Имеет смысл браться за LSTM и GRU только если есть длинные плотные ряды и вычислительные ресурсы под это дело.

Архитектуры на основе механизма внимания (Transformer, Temporal Fusion Transformer) – наиболее современное направление. TFT, предложенный Lim et al. [10], явно разделяет статические метаданные (характеристики направления) и временные паттерны, обеспечивая интерпретируемость через веса внимания. Высокие вычислительные требования и сложность настройки пока ограничивают их распространение в производственных системах.

Сравнение программных инструментов

Выбор инструментов влияет на практическую применимость системы. Рисунок 2 сравнивает ключевые инструменты по шести параметрам: наличие открытого исходного кода, скорость, масштабируемость, качество решения, интерпретируемость и области применения в логистике.

Сравнение инструментов для оптимизации и прогнозирования в логистике

Инструмент	Открытый исходный код	Скорость вычислений	Масштабируемость	Качество решения	Интерпретируемость	Применение в логистике
IBM CPLEX	Нет	Низкая	Средняя	Высокое	Высокая	MILP, VRP
OR-Tools (Min-Cost Flow)	Да (Apache 2.0)	Высокая	Высокая	Оптимальное	Высокая	Сетевые потоки
LightGBM	Да (MIT)	Высокая	Высокая	Высокое	Высокая	Прогноз спроса, времени в пути
LSTM / GRU (нейросети)	Да (PyTorch)	Низкая	Средняя	Среднее	Низкая	Временные ряды, последовательности
Prophet	Да (MIT)	Средняя	Низкая	Среднее	Высокая	Сезонные ряды, один ряд за раз
XGBoost	Да (Apache 2.0)	Средняя	Высокая	Высокое	Высокая	Прогноз спроса, табличные данные

Рисунок 2 – Сравнение инструментов оптимизации и прогнозирования в задачах логистики

OR-Tools от Google – открытая библиотека (Apache 2.0) для широкого класса задач комбинаторной оптимизации: MCF, CP-SAT и VRP-солвер. На задачах с сотнями узлов решение MCF занимает менее секунды. На сложных MILP-задачах уступает CPLEX, однако для постановки MCF разница несущественна – оба находят точный оптимум.

IBM CPLEX – коммерческий решатель для MILP и LP. Даёт наилучшее качество на задачах смешанной целочисленности, но требует лицензионных затрат (порядка нескольких миллионов рублей в год для корпоративного использования) и серверной инфраструктуры с большим объёмом RAM.

LightGBM показывает лучшее сочетание характеристик для прогнозирования спроса в логистике: высокая скорость обучения, масштабируемость на миллионы строк, нативная поддержка категориальных признаков и интерпретируемость через механизм важности признаков. XGBoost занимает ту же нишу, уступая LightGBM в скорости на больших датасетах.

Интеграция прогнозирования и оптимизации

Наиболее перспективное направление в динамической маршрутизации – объединение прогностических моделей и алгоритмов оптимизации в единый вычислительный конвейер. Исторически эти задачи решались независимо: аналитики строили прогноз спроса и передавали его плановикам для составления маршрутной схемы. При таком двухэтапном подходе прогнозные ошибки не учитываются оптимизатором, а обратная связь от оптимизационного решения не влияет на прогнозную модель.

Ряд работ предлагает архитектуры, в которых прогностические модели непосредственно задают параметры оптимизационной задачи. В работе [11] авторы показывают, что подача прогнозных оценок спроса как параметров задачи MCF – а не фиксированных коэффициентов – снижает суммарный транспортный километраж на 15–30% по сравнению с историческими средними. Аналогичные результаты получены в авиационной [12] и железнодорожной логистике.

Практический пример такого конвейера – система динамического планирования маршрутов для магистральной логистики крупного российского маркетплейса. Система состоит из четырёх блоков: подготовка данных (IQR-фильтрация, удаление аномалий), прогноз суточного спроса по направлениям (LightGBM, MAE-функция потерь, 12 признаков), прогноз времени в пути (LightGBM, Huber-loss) и оптимизация маршрутной схемы (MCF, OR-Tools SSP). На данных 2024 года (1,3 млн записей, 209 складских узлов) конвейер снизил суммарную взвешенную стоимость маршрутной схемы на 67% по сравнению с прямыми маршрутами без оптимизации. Показательно другое: MCF с историческими средними вместо прогнозов даёт результат на 23% хуже, чем решение без оптимизации вообще. Это наглядно показывает, что качество входных данных здесь не менее важно, чем сам алгоритм.

Система полностью построена на открытых инструментах (Python, LightGBM, OR-Tools, pandas) и работает на обычном персональном компьютере – в отличие от промышленных решений на базе коммерческих MILP-солверов.

Заключение

В статье систематизированы методы и инструменты динамической маршрутизации грузовых потоков. Ни один из рассмотренных классов не универсален: точные методы гарантируют оптимум, но медленные; эвристики масштабируются, но без гарантий качества; методы машинного обучения адаптируются к данным, однако это инструмент прогнозирования, а не оптимизации.

Объединение прогностических моделей и оптимизационных алгоритмов в единый конвейер остаётся наиболее продуктивным направлением. Задача минимального стоимостного потока здесь занимает особое место: решается за полиномиальное время и хорошо соответствует структуре задач магистральной логистики – в отличие от VRP, требующей учёта транспортных средств и расписаний. MCF совместно с прогностическими моделями на основе градиентного бустинга позволяет строить адаптивные маршрутные схемы на открытом стеке без коммерческих лицензий.

Открытые вопросы для дальнейших исследований: интеграция данных о дорожной обстановке в реальном времени, механизмы онлайн-обучения для адаптации моделей без полного переобучения, а также постановка Multi-Commodity Flow для задач с разнородными типами грузов.

Литература

1. Toth P., Vigo D. (Eds.) Vehicle Routing: Problems, Methods, and Applications. 2nd ed. – Philadelphia: SIAM, 2014. – 481 p.
2. Achterberg T., Wunderling R. Mixed Integer Programming: Analyzing 12 Years of Progress // Facets of Combinatorial Optimization. – Berlin: Springer, 2013. – P. 449–481.

3. Ahuja R.K., Magnanti T.L., Orlin J.B. Network Flows: Theory, Algorithms, and Applications. – New Jersey: Prentice Hall, 1993. – 846 p.
4. Goldberg D.E. Genetic Algorithms in Search, Optimization, and Machine Learning. – Reading: Addison-Wesley, 1989. – 412 p.
5. Dorigo M., Stützle T. Ant Colony Optimization. – Cambridge: MIT Press, 2004. – 305 p.
6. Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M. Time Series Analysis: Forecasting and Control. 5th ed. – Hoboken: Wiley, 2015. – 712 p.
7. Taylor S.J., Letham B. Forecasting at Scale // The American Statistician. – 2018. – Vol. 72, No. 1. – P. 37–45.
8. Ke G., Meng Q., Finley T. [et al.]. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 3146–3154.
9. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735–1780.
10. Lim B., Arik S.Ö., Loeff N., Pfister T. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting // International Journal of Forecasting. – 2021. – Vol. 37, No. 4. – P. 1748–1764.
11. Perrone G., Vasiliev I., Nikulin V. Machine Learning for Demand Forecasting in the Logistics Sector // Expert Systems with Applications. – 2023. – Vol. 213. – Article 119101.
12. Bertsimas D., Gupta V., Kallus N. Prescriptive Analytics for Flight Delay Reduction // Operations Research. – 2019. – Vol. 67, No. 3. – P. 655–675.

Literature

1. Toth P., Vigo D. (Eds.) Vehicle Routing: Problems, Methods, and Applications. 2nd ed. – Philadelphia: SIAM, 2014. – 481 p.
2. Achterberg T., Wunderling R. Mixed Integer Programming: Analyzing 12 Years of Progress // Facets of Combinatorial Optimization. – Berlin: Springer, 2013. – P. 449–481.

3. Ahuja R.K., Magnanti T.L., Orlin J.B. Network Flows: Theory, Algorithms, and Applications. – New Jersey: Prentice Hall, 1993. – 846 p.
4. Goldberg D.E. Genetic Algorithms in Search, Optimization, and Machine Learning. – Reading: Addison-Wesley, 1989. – 412 p.
5. Dorigo M., Stützle T. Ant Colony Optimization. – Cambridge: MIT Press, 2004. – 305 p.
6. Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M. Time Series Analysis: Forecasting and Control. 5th ed. – Hoboken: Wiley, 2015. – 712 p.
7. Taylor S.J., Letham B. Forecasting at Scale // The American Statistician. – 2018. – Vol. 72, No. 1. – P. 37–45.
8. Ke G., Meng Q., Finley T. [et al.]. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 3146–3154.
9. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735–1780.
10. Lim B., Arik S.Ö., Loeff N., Pfister T. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting // International Journal of Forecasting. – 2021. – Vol. 37, No. 4. – P. 1748–1764.
11. Perrone G., Vasiliev I., Nikulin V. Machine Learning for Demand Forecasting in the Logistics Sector // Expert Systems with Applications. – 2023. – Vol. 213. – Article 119101.
12. Bertsimas D., Gupta V., Kallus N. Prescriptive Analytics for Flight Delay Reduction // Operations Research. – 2019. – Vol. 67, No. 3. – P. 655–675.