Следом за потоком

Расширенное исследование MOEX · 20 октября 2026

Что уже можно сказать

Идея следовать за устойчивым потоком заслуживает проверки. Но признаков, что нынешние детекторы уже дают прибыль после издержек, нет.

Новая выборка содержит 70 500 биржевых сделок по 83 инструментам: 65 акций и 18 фьючерсов. Она охватывает 20 октября 2026 года, 19:34–20:28 МСК. Это один вечер, а не вся Московская биржа и не независимое продолжение предыдущей выборки. Часть интервала пересекается с первым исследованием.

На отложенной по времени части простое сглаживание угадывает направление следующей группы в 65,17% случаев, байесовская модель в 64,86%. Это направление сделок, не рост цены. Среднее движение цены через две минуты после 21 подходящего сигнала составило +2,19 б.п. до издержек. Принятое допущение по комиссии и проскальзыванию составляет 12 б.п. за круг, ещё до спреда.

Запущен отдельный непрерывный сборщик и четыре бумажных варианта. Они показывают, когда найден поток, когда удалось войти, почему произошёл выход и сколько осталось после издержек. Основной виртуальный портфель не объединяется с этими счетами.

Открыть живой эксперимент · Скачать этот отчёт · Первое исследование

Дополнение к первому исследованию. 30 первичных источников в общей библиографии. Факты нашей выборки получены собственными скриптами; внешние исследования не подтверждают прибыль нашей системы.

Что именно мы пытаемся заметить

Рабочая гипотеза: крупное исполнение оставляет устойчивый направленный поток. Если заметить начало достаточно рано, движение иногда может продолжиться после нашего входа. Задача системы состоит из четырёх отдельных проверок: возник ли необычный поток, продолжится ли он, отвечает ли цена и перекрывает ли оставшееся движение издержки.

Нельзя считать каждую серию одним роботом. Несколько участников могут покупать одновременно; одна заявка может разбиться на множество исполнений. На MOEX существуют штатные механизмы TWAP и айсберг-заявок. Их существование не позволяет приписать конкретную серию конкретному алгоритму. MOEX: запуск TWAP · MOEX: айсберг-заявки.

Наблюдение	Что допускает	Чего не доказывает
Повтор объёма и интервала	Гипотеза дробления исполнения	Один владелец или конкретный робот
Усиление одной стороны	Изменение режима потока	Будущую прибыль
Поток есть, цена стоит	Возможное поглощение ликвидностью	Достоверный айсберг
Котировка быстро исчезла	Изменение видимой ликвидности	Манипулятивное намерение
Синхронность связанных контрактов	Кандидат на совместную динамику	Общий аккаунт или арбитражёра

Исследование режимов Tsaknaki, Lillo и Mazzarisi показывает полезность онлайн-моделей потока на NASDAQ. Авторы учитывают зависимости внутри режима. Наша текущая Beta-Bernoulli модель проще: предполагает условную независимость внутри режима и не воспроизводит их полную score-driven модель. Первичная статья, 2024.

Данные и проверка без будущего в признаках

Из обменных сделок сформированы 33 408 групп. Сделки одного инструмента и направления в одном окне 100 мс объединены, объёмы суммированы. Это уменьшает влияние дробных исполнений, но не восстанавливает исходную заявку или участника.

Первые 70% групп каждого инструмента служат прогревом. На последних 30% прогноз вычисляется до добавления текущего наблюдения в модель. На малых выборках требуется минимум 60 групп прогрева. Из оценки исключены 303 наблюдения, соседние с противоположным направлением в том же временном окне. Получено 9 581 оценённое направление.

Разделение хронологическое, однако исследование ретроспективное. Файл фиксации параметров v0.1 создан после периода исторической выборки, перед живым запуском. Быстрый EWMA и марковский контроль добавлены во время анализа. Поэтому эту таблицу нельзя называть заранее зарегистрированным независимым подтверждением. Параметры живых вариантов сохранены отдельно с хешем исходного кода.

В историческом тесте нет задержки доставки и нет исполняемого стакана в момент входа. В живом сборщике есть буфер 500 мс после окончания группы плюс минимум 250 мс ожидания исполнения и следующей котировки. Историческая точность не измеряет доступное для торговли опережение.

Цена через 30 или 120 секунд определяется первой сделкой после горизонта с допуском 2 секунды. При отсутствии такой сделки результат пропускается. Это прокси движения, не симуляция исполнения. Соседние сигналы и инструменты зависимы; доверительные интервалы при предположении независимых сделок здесь были бы вводящими в заблуждение.

Артефакты: collection.json, raw/*.json.gz, replay.py, replay-summary*.json, instrument-comparison*.csv. Исходный engine-v0.1-frozen.py совпадает с сохранённым SHA-256. Новая логика выхода v0.2 не участвовала в историческом сравнении.

Сложность модели пока не окупилась



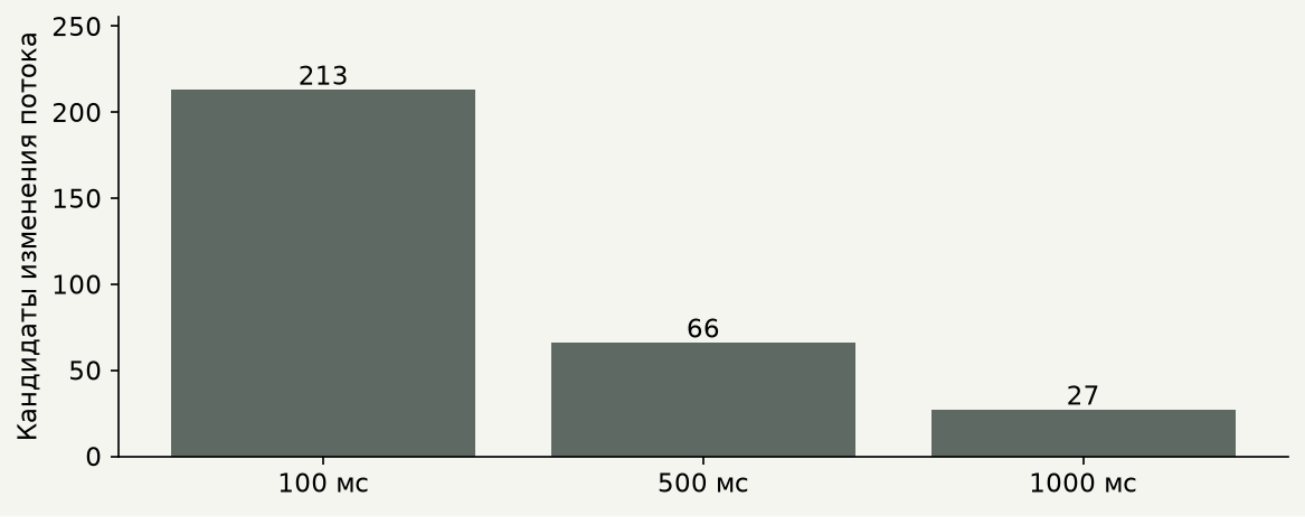
Модель	Точность, %	Brier ↓	Log loss ↓
Постоянная 50%	50,00	0,250000	0,693147
EWMA 0,02	61,07	0,233184	0,658652
EWMA 0,10	65,17	0,222123	0,638490
Марковская	63,80	0,227581	0,647199
BOCPD	64,86	0,221656	0,634420

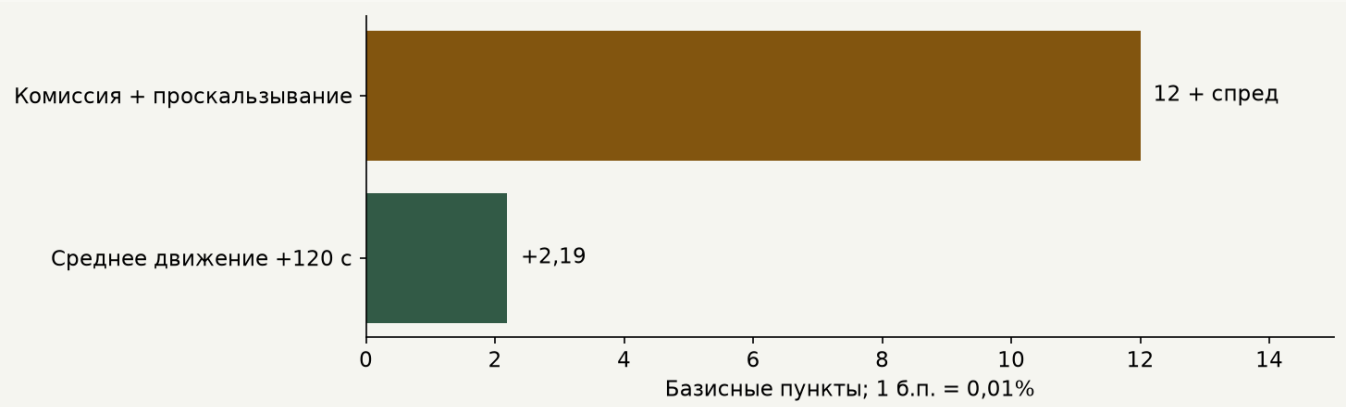
BOCPD чуть лучше по вероятностной ошибке Brier: 0,221656 против 0,222123 у быстрого EWMA. По доле верных направлений быстрый EWMA выше. Размер преимущества и единственный короткий период не дают оснований объявлять алгоритмический прорыв.

Реализация BOCPD хранит до 64 вариантов длины режима, prior Beta(2,2), вероятность сброса 1/80 на группу. CUSUM накапливает свидетельство отклонения от EWMA с шагом 0,02. Порог CUSUM 5, направленная вероятность не ниже 0,65, минимум 6 групп за 15 секунд и минимум 65% объёма в одну сторону. Это эвристические параметры, а не оценённая вероятность обнаружения робота.

Следующее обоснованное усложнение: моделировать зависимость последовательных направлений внутри режима и условный отклик цены. Добавлять сложность следует только после сравнения на новых днях с теми же простыми контролями.

Устойчивость и стоимость входа





Один и тот же вечер даёт 213 кандидатов изменения потока при группировке 100 мс, 66 при 500 мс и 27 при 1 секунде. Частота сильнее зависит от способа агрегации, чем хотелось бы. Эти количества относятся ко всей выборке, не только к отложенной части, и не являются числом роботов.

Для отложенной части доступен 21 ценовой исход через 120 секунд, среднее +2,19 б.п., медиана +1,87 б.п. Через 30 секунд: 18 исходов, среднее +0,63 б.п. Для детектора четырёх повторов всего по одному исходу: выводов делать нельзя.

12 б.п. издержек на графике включают принятую комиссию 5 б.п. и проскальзывание 1 б.п. на каждой стороне. Реальный спред добавляется отдельно. Это исследовательское допущение, не проверенный тариф пользователя. При другом тарифе граница окупаемости изменится, но её нельзя заменять нулевыми издержками.

Четыре работающих бумажных варианта

Вариант	Вход	Отличие
4 повтора	4 близких интервала, одинаковый объём >1 лота	Базовый детектор регулярного исполнения
Изменение потока	CUSUM + вероятность направления + объём	Не ждёт четырёх одинаковых порций
Поток + цена	Тот же поток + отклик цены и дисбаланс L1	Отсекает часть поглощаемого потока
С ранним выходом	Такой же вход, как у «Поток + цена»	Выходит при устойчивой потере отклика

Подтверждение цены: направленное изменение середины стакана за 5 секунд выше максимума из 0,5 б.п. и четверти текущего спреда. Дисбаланс объёма лучшего уровня должен согласовываться с направлением. После сигнала подтверждение ожидается до 10 секунд.

Дополнительный ранний выход: после 10 секунд удержания поток за последние 5 секунд содержит минимум 3 группы и минимум 65% объёма в сторону позиции, но направленное движение цены не выше 0,25 б.п. Если это сохраняется 3 секунды, позиция закрывается. Это проверяемая гипотеза о потере отклика, не знание о завершении чужой программы.

Общие выходы: движение против входа 25 б.п., цель 40 б.п., 120 секунд удержания либо 15 секунд без сделок нужной стороны после первых 10 секунд позиции. Порог стопа измеряется до комиссий; итоговый убыток может оказаться больше. При отсутствии свежего стакана выход ждёт данные.

Один лот на вход. На каждый вариант до 5 позиций, до 100 тыс. ₽ на позицию и до 500 тыс. ₽ суммарного номинала. После реализованной потери 30 тыс. ₽ новые входы блокируются. Каждый счёт имеет независимую начальную базу 1 млн ₽, которая не суммируется с основным портфелем. Малый размер здесь специально фиксирован для сравнения сигналов.

Ранний выход добавлен позже трёх первоначальных вариантов. Его следует сравнивать только на совпадающих signal_id и одинаковых ценах входа. Общие накопленные суммы за разные периоды несопоставимы. Ограничения позиции могут давать непарные входы, их нужно выделять отдельно.

Первые наблюдения живого запуска

Контрольный снимок: 20.10.2026 21:17:16 МСК. Это зафиксированная точка исследования; текущие значения находятся на живой странице.

Вариант	Закрыто	Прибыльных	Реализовано, ₽	Открыто
4 повтора	10	0	-360,40	0
Изменение потока	36	0	-990,10	1
Поток + цена	19	0	-474,24	0
С ранним выходом	0	0	0,00	0

Первые три варианта уже дали отрицательные результаты. Их история не сброшена при обновлении программы. У новой версии нет основания наследовать прибыльность или статистику старых вариантов. Отрицательные результаты являются полезным ограничением гипотезы, а не поводом увеличить плечо.

Выполнения виртуальные: по актуальному bid/ask, с неблагоприятным проскальзыванием и комиссией. Для фьючерсов учитывается стоимость шага, цены округляются неблагоприятно по доступному шагу. Очередь заявок, рыночное воздействие, частичное исполнение, доступность шорта и фактическое обеспечение брокера не моделируются. Для акций эффективное проскальзывание может давать цену между тиками.

Часть самых первых сигналов не могла исполниться до исправления начальной загрузки статусов торгов. Эти отказы сохранены. При разрыве потока признаки и незавершённые сигналы очищаются, позиции и накопленные итоги восстанавливаются из контрольной точки.

Что работает постоянно и что видит ИИ

Отдельный сервис наблюдает 83 инструмента через обезличенные биржевые сделки и стаканы глубиной 10. Используются только методы рыночных данных. В коде нет отправки приказов брокеру. Работа продолжается без открытой вкладки; systemd запускает сервис после перезагрузки и восстанавливает его после ошибки.

Локальный алгоритм реагирует на события потока. Нейросеть не вызывается для каждого тика. Ежедневный разбор Gemini получает количество исходов, прибыль и убытки, причины выходов, свежесть наблюдений и параметры симуляции. Сохраняется бюджет 19 вызовов основной работы и 1 разбора в сутки. Подключение нового контекста само по себе не требует дополнительного запроса.

На странице: линии результата, сравнение четырёх вариантов, открытые позиции с временем удержания, последние серии, причины закрытий и поиск по инструментам. При потере связи статус меняется, старые данные не выдаются за живые. На графике нет придуманных значений нового варианта до его запуска.

Архив ограничен 320 МиБ: 192 для сырых данных, 96 для признаков раз в минуту, 32 для событий. При заполнении удаляются старые файлы соответствующей категории. При остатке диска менее 512 МиБ запись архива приостанавливается. Поэтому полный многонедельный архив тиков не обещается. Контрольная точка хранит накопленные итоги, открытые позиции, последние 200 закрытий и около часа графика.

Ключи остаются в серверном хранилище с ограниченными правами. Публичная страница получает только статистику. Доступ к исследованию через основной экран MOEX Lab.

Как превратить идею в сильную проверку

Направление	Следующий эксперимент	Критерий пользы
Раннее обнаружение	Сравнить задержку CUSUM и повторов на общем событии	Меньшая задержка без роста ложных тревог
Продолжение исполнения	Оценить вероятность паузы по возрасту и темпу режима	Калибровка на новых днях лучше EWMA
Оставшееся движение	Прогнозировать изменение исполняемой цены после сигнала	Доходность после спреда и комиссии
Поглощение	Сравнить одинаковые входы с обычным и ранним выходом	Положительная парная разность результатов
Связанные инструменты	Проверить лаг крупных и мини-контрактов, акций и фьючерсов	Результат после задержки, стоимости ног и базиса
Адаптация	Отдельно калибровать ликвидность и время торгов	Устойчивость на будущих временных блоках

Это план дальнейших гипотез, кроме уже работающего простого раннего выхода. Полная модель скрытого исполнения, идентификация участников и прибыльный межконтрактный арбитраж не заявлены как реализованные.

Перед выбором победителя следует зафиксировать горизонты, комиссии и критерий, собрать новые торговые дни, отделить дни настройки от дней оценки, исключить пересекающиеся исходы либо группировать ошибки по дню и инструменту. Не выбирать параметры по лучшему результату единственного вечера.

Практический фильтр для будущей версии: вход допустим, только если нижняя оценка оставшегося движения перекрывает полную стоимость входа и выхода с запасом. Сейчас такой откалиброванной модели нет; её нельзя заменить процентом «уверенности» языковой модели.

До подтверждения положительного результата после издержек увеличение размера сделок лишь масштабирует непроверенный эффект. Первое доказательство должно быть получено на новых данных и сопоставимых виртуальных входах.

Покрытие инструментов

65 акций и 18 фьючерсов. Есть золото, серебро, платина, палладий, другие металлы, нефть, газ, валютные и индексные контракты. Состав фиксирован для этой версии; автоматическая замена истекающего контракта пока не реализована.

Класс	Инструменты
Акции	CHMF, GAZP, GMKN, LKOH, MOEX, MTSS, NVTK, ROSN, SBER, TATN, AFKS, AFLT, ALRS, BSPB, CBOM, ENPG, FEES, FLOT, HYDR, IRAO, MAGN, MGNT, MSNG, MTLR, MTLRP, NLMK, OGKB, PHOR, PIKK, PLZL, RASP, RTKM, RTKMP, RUAL, SBERP, SNGS, SNGSP, TATNP, TRNFP, UPRO, VKCO, YDEX, T, OZON, HEAD, POSI, ASTR, SOFL, SELG, SGZH, SFIN, RENI, LEAS, UGLD, GCHE, BELU, LSRG, ETLN, DIAS, DELI, SVAV, MSRS, MSST, LSNG, LSNGP
Фьючерсы	GLZ6, S1Z6, LTZ6, LDZ6, CEZ6, BMX6, NGV6, UMZ6, CRZ6, MMZ6, GDZ6, GNZ6, SVZ6, PTZ6, PDZ6, BRX6, NRV6, SiZ6

Наиболее активные в исторической выборке

Инструмент	Сделок	Групп	Проверок прогноза	Сигналов потока
GAZP	5632	2509	727	9
BRX6	4980	2336	678	7
MMZ6	3429	1343	373	8
SBER	3290	2043	599	6
BMX6	2895	1409	418	8
NVTK	2372	962	277	2
T	2249	1185	350	5
ROSN	2247	1097	327	8
NGV6	1822	858	251	2
SiZ6	1768	752	216	7
RUAL	1403	481	140	6
LKOH	1385	688	201	7
AFKS	1376	504	141	5
IRAO	1327	647	187	7
YDEX	1221	598	174	2

Полная таблица всех инструментов находится в instrument-comparison.csv. Малоликвидные инструменты могут не набрать необходимый прогрев и потому не участвовать в сравнении точности.

Первичные источники

Библиография включает первый отчёт и четыре дополнительных источника. Наличие ссылки не означает воспроизведение всей методики статьи. Описание применимости приведено рядом.

1. TWAP на MOEX: запуск 22 июня 2026. MOEX.
   Факт запуска TWAP на TQBR; не идентификация конкретной серии.
2. Параметры алгоритма TWAP. MOEX.
   Дробление, случайные отклонения времени и объёма; описание механизма.
3. Айсберг-заявки срочного рынка. MOEX.
   Обновление видимой части и потеря приоритета новой порцией; не доказательство айсберга по агрегированному стакану.
4. Программы маркет-мейкинга срочного рынка. MOEX.
   Наличие программ и классов контрактов; не атрибуция обезличенных сделок.
5. Полный журнал заявок Full_orders_log. MOEX.
   Анонимные события и дополнительная стоимость доступа, не цена всей инфраструктуры.
6. SIMBA SPECTRA. MOEX.
   Прямой поток для колокации и полный журнал с оговорёнными исключениями.
7. Тарифы срочного рынка. MOEX.
   Различие комиссий maker/taker; брокерские издержки отдельно.
8. Лимиты T-Invest API. Т-Банк.
   Стакан не чаще 100 мс, лимиты соединений и подписок.
9. MarketDataService: схемы и методы. Т-Банк.
   TradeSource, обезличенные сделки за последний час, поля Trade.
10. FAQ по рыночным данным. Т-Банк.
    Лоты, цена и стоимость шага; единицы нельзя смешивать.
11. Рыночные данные: обзор сервиса. Т-Банк.
    Потоковые и исторические данные, особенности свечей.
12. Высокочастотная торговля, первая часть. Банк России.
    Российский рынок, идентификация HFT по более подробным данным; не доля HFT в 2026.
13. Высокочастотная торговля, вторая часть. Банк России.
    Влияние разных стратегий на волатильность; историческая выборка.
14. Резюме исследования волатильности. Банк России.
    Нет единого систематического направленного эффекта всей группы HFT в исследованной выборке.
15. The price impact of order book events. Cont, Kukanov, Stoikov.
    Связь OFI и изменения цены на 50 акциях США; не готовый прогноз прибыли.
16. Queue Imbalance as a One-Tick-Ahead Price Predictor. Gould, Bonart.
    Прогноз следующего движения mid-price на Nasdaq; не исполнимая доходность MOEX.
17. CME Iceberg Order Detection and Prediction. Zotikov, Antonov.
    Детекция по событиям и идентификаторам заявок CME; протокол не равен нашему L1.
18. Why is equity order flow so persistent?. Tóth, Palit, Lillo, Farmer.
    Роль дробления заявок в данных LSE с кодами участников; перенос ограничен.
19. High-frequency trading in a limit order book. Avellaneda, Stoikov.
    Модель управления запасом маркет-мейкера; не гарантированная стратегия.
20. High Frequency Lead/lag Relationships. Huth, Abergel.
    Прогноз направления может терять прибыль из-за спреда.
21. The Deflated Sharpe Ratio. Bailey, López de Prado.
    Поправка на множественный поиск и ненормальность доходностей.
22. Hawkes processes in finance. Bacry, Mastromatteo, Muzy.
    Модели кластеров событий; кластеризация сама по себе не раскрывает участника.
23. Learning the Spoofability of Limit Order Books. Fabre, Challet.
    L3 крипторынка и положение заявок; не классификатор манипуляций MOEX по L1.

24. Спецификации фьючерсов на драгметаллы. MOEX.
Реестр действующих спецификаций; сами doc-приложения отдельно не разбирались.

25. КИД расчётных фьючерсов на драгметаллы. MOEX.
Единицы контрактов и валюты котировки; при реализации сверять действующую спецификацию.

26. Параметры поставочных фьючерсов на металлы. MOEX.
Другой базис поставки и единицы рублёвого золота; контракты не взаимозаменяемы автоматически.

27. Online Learning of Order Flow and Market Impact with Bayesian Change-Point Detection Methods. Tsaknaki, Lillo, Mazzarisi.
Модель режимов на NASDAQ; не доказательство прибыли на MOEX.

28. Bayesian Online Changepoint Detection. Adams, MacKay.
Основа распределения длины режима.

29. Statistical identification with hidden Markov models of large order splitting strategies in an equity market. Vaglica, Lillo, Mantegna.
Методы сегментации с данными об участниках; такие идентификаторы у нас отсутствуют.

30. Optimal Execution with Dynamic Order Flow Imbalance. Bechler, Ludkovski.
Динамика дисбаланса и исполнение; не идентификация чужих роботов.