Информационные процессы, Том 18, № 1, 2018, стр. 55–71
© 2018 В.В. Вьюгин, В.Г. Трунов.

ТЕОРИЯ И МЕТОДЫ ОБРАБОТКИ ИНФОРМАЦИИ

Отслеживание оптимальной последовательности предсказательных стратегий[1]

В.В. Вьюгин, В.Г. Трунов

Институт проблем передачи информации им. А.А. Харкевича РАН, Москва, Россия
Поступила в редколлегию 19.03.2018

Аннотация—В рамках теории предсказания (принятия решений) с экспертами в режиме онлайн предложен адаптивный алгоритм, который агрегирует решения экспертных стратегий и несет потери, не превосходящие (с точностью до некоторой величины, называемой регретом) потери наилучшей комбинации экспертов, произвольным образом распределенных по интервалу прогнозирования. Алгоритм развивает метод Mixing Past Posteriors и алгоритм экспоненциального взвешивания экспертных решений AdaHedge, использующий адаптивный параметр обучения. Получена оценка регрета предложенного алгоритма. В рамках используемого подхода не делается никаких предположений о природе источника исходных данных и об ограниченности потерь экспертов. Приведены результаты численных экспериментов по смешиванию экспертных решений с помощью предложенного алгоритма в условиях высокой волатильности потерь экспертов.

Ключевые слова: алгоритмы распределения потерь в режиме онлайн, предсказания с использованием экспертных стратегий, схемы смешивания апостериорных распределений экспертов, адаптивный параметр обучения.

1. ВВЕДЕНИЕ

Рассматривается задача принятия оптимальных решений в режиме онлайн (Decision Theoretic Online Learning). Имеется набор методов (экспертов, алгоритмов), которые на каждом шаге принимают решения (или выдают прогнозы) и несут потери вследствие своих решений (см. [2–7] и др.).

На каждом раунде (шаге) игры $t = 1, 2, \ldots, T$ алгоритм принимает некоторое решение – выдает вектор распределения потерь $w_t = (w_{1,t}, \ldots, w_{N,t})$, где $\sum_{i=1}^N w_{i,t} = 1$ и $w_{i,t} \geq 0$ для всех i. После этого свои решения принимают эксперты $i = 1, \ldots, N$, в результате чего они несут потери l_t^i, $i = 1, \ldots, N$. Потери агрегирующего алгоритма равны $h_t = \sum_{i=1}^N w_{i,t} l_t^i$. В процессе игры накапливаются кумулятивные потери каждого эксперта $L_T^i = \sum_{t=1}^T l_t^i$, а также кумулятивные потери алгоритма $H_T = \sum_{t=1}^T h_t$.

Цель агрегирующего алгоритма – принять оптимальное решение, используя решения (прогнозы) экспертов так, чтобы его потери $H_T = \sum_{t=1}^T h_t$ не превосходили потери любого эксперта i плюс небольшие потери в процессе обучения – регрет $R_T^i = H_T - L_T^i$. Алгоритм должен принимать такие решения, чтобы минимизировать величину $R_T = \max_{1 \leq i \leq N} R_T^i$ – минимаксный регрет.

[1] Работа выполнена при финансовой поддержке гранта 14-50-00150 Российского научного фонда. Предварительный вариант работы представлен в трудах конференции [1].

В [3] предложен алгоритм экспоненциального взвешивания экспертных решений Hedge, который на каждом шаге адаптирует веса экспертов $i = 1, \ldots, N$ согласно правилу

$$w_{i,t+1} = \frac{w_{i,t} e^{-\eta l_t^i}}{\sum\limits_{j=1}^{N} w_{j,t} e^{-\eta l_t^j}},$$

где η – параметр обучения (learning rate). Этот параметр влияет на оценку регрета и стабильность работы алгоритма.

Ряд других алгоритмов для решения этой задачи приведен в [6]. Основные алгоритмы агрегирования экспертных решений предполагают ограниченность потерь экспертов на каждом шаге, например, предполагается $l_t^i \in [0, 1]$ для всех i и t. Для этого случая получены верхние оценки регрета, которые служат гарантией эффективности рассматриваемых алгоритмов. Другая проблема – это проблема настройки параметра обучения η.

Предложенная в [8] версия алгоритма Hedge достаточно эффективно решает эти задачи. Алгоритм AdaHedge использует на каждом шаге t адаптивный (переменный) параметр обучения $\eta = \eta_t$, который определяется в зависимости от регрета, накопленного на предыдущих шагах $t' < t$ работы алгоритма.

При получении оценок регрета этого алгоритма не делается никаких предположений об ограниченности (и знаке) потерь экспертов на каждом шаге. В этом случае получена гарантийная оценка эффективности работы алгоритма в условиях высокой волатильности потерь экспертов.

Пусть $l_t^- = \min_i l_t^i$ и $l_t^+ = \max_i l_t^i$ – наименьшие и наибольшие потери среди экспертов на шаге t. Определим $L_T^+ = \sum_{t=1}^T l_t^+$ и $L_T^- = \sum_{t=1}^T l_t^-$. Пусть также $s_t = l_t^+ - l_t^-$ и $S_T = \max\{s_1, \ldots, s_T\}$. Также $L_T^* = \min_{1 \le i \le N} L_T^i$ – потери наилучшего эксперта на шагах $1 \le t \le T$.

Алгоритм AdaHedge позволяет получить оценку для регрета алгоритма при отсутствии ограничений на потери экспертов:

$$R_T \le 2\sqrt{S_T \frac{(L_T^* - L_T^-)(L_T^+ - L_T^*)}{L_T^+ - L_T^-} \ln N} + \left(\frac{16}{3} \ln N + 2\right) S_T. \tag{1}$$

В случае ограниченных пошаговых потерь экспертов данная оценка превращается в стандартную по порядку оценку $O(\sqrt{T \ln N})$.

Другой, более общий, подход к агрегированию экспертных стратегий, предложенный в [9], состоит в следующем: серия шагов, на которых делаются предсказания, делится на $k+1$ сегментов. Каждому сегменту ставится в соответствие свой эксперт; последовательность сегментов и соответствующих экспертов называется составным экспертом. Цель алгоритма изменяется – теперь он должен предсказывать так, чтобы быть не хуже каждого составного эксперта. Соответственно модифицируется понятие регрета алгоритма – теперь это разность между потерями алгоритма и потерями наилучшей последовательности экспертов. Данное изменение позволяет точнее моделировать условия реальной жизни, когда природа исходов может меняться со временем и разные эксперты могут предсказывать с разной степенью успешности в зависимости от текущего тренда. Соответствующий алгоритм называется Fixed Share.

В статье будет использоваться несколько более общий метод Mixing Past Posteriors (MPP), предложенный в [10]. Алгоритм Fixed Share является частным случаем метода MPP. Формальная постановка задачи и соответствующий алгоритм будут приведены в разделе 2.

Рассматриваемые в публикациях версии алгоритма MPP для соответствующих оценок их регрета используют постоянный параметр обучения, тогда как алгоритм AdaHedge основан

на использовании адаптивного (переменного) параметра обучения. Для анализа работы совмещенной версии этих алгоритмов необходимо получить оценку регрета алгоритма MPP с переменным параметром обучения.

В данной статье представлен алгоритм, который позволяет применить метод MPP к алгоритму AdaHedge и получить верхнюю оценку для величины регрета порядка $O(k \ln T R_T)$, где R_T – регрет алгоритма AdaHedge, $k+1$ – число сегментов составного эксперта, с которым производится сравнение, а в правой части (1) L_T^* обозначает потери наилучшего эксперта, составленного из $k+1$ сегментов. Данная оценка является несколько более слабой, чем (1), однако преимущество предлагаемого в статье алгоритма заключается в том, что он приближает свои потери к меньшей величине – потерям наилучшего составного эксперта.

Данный алгоритм обладает преимуществами алгоритма из [8] и алгоритма MPP. В работе представлена модифицированная версия метода MPP, использующая этот алгоритм и включающая в себя переменные параметры. Для этого случая получена верхняя оценка регрета модифицированного алгоритма.

Алгоритм имеет оценку времени работы $O(T)$ в отличие от аналогичного алгоритма, предложенного в [11], который основан на другой вычислительной модели и имеет оценку времени работы $O(T^2)$. При этом оба алгоритма имеют приблизительно одинаковые верхние оценки регрета.

В рамках данного подхода не делается никаких стохастических предположений об источнике исходных данных и об ограниченности и знаке потерь экспертов.

В разделе 2 приведена модифицированная версия алгоритма MPP, в которой предложена новая процедура перераспределения весов. Получена верхняя оценка регрета этого алгоритма. В разделе 3 приведено доказательство основного результата статьи – теоремы 2.

В экспериментальном разделе 4 приведены результаты численных расчетов, проведенных с помощью предложенного алгоритма на искусственных и реальных данных.

Первый эксперимент, проведенный на искусственных данных, демонстрирует работу предложенного алгоритма в условиях, когда потери экспертов на одном шаге не ограничены и могут быть как положительными так и отрицательными.

Экспертные стратегии во втором эксперименте – это стратегии, которые покупают и продают акции в режиме онлайн. В этом случае потери заменяются на выигрыши (эти величины противоположны по знаку). В этой работе, в экспериментах, используются только динамические ряды выигрышей этих стратегий. Детально эти стратегии изучаются в [12, 13].

Адаптивный алгоритм из раздела 2 динамически перераспределяет финансовые средства между этими стратегиями в зависимости от их доходности. Отметим, что временные ряды выигрышей этих стратегий обладают высокой степенью волатильности.

Результаты экспериментов показали преимущество алгоритма MPP с переменным параметром обучения по сравнению с алгоритмом AdaHedge.

2. АЛГОРИТМ MPP С АДАПТИВНЫМ ПАРАМЕТРОМ ОБУЧЕНИЯ

Рассмотрим метод смешивания апостериорных распределений экспертов с использованием схем смешивания общего вида Mixing Past Posteriors (MPP), предложенный в [10].

Предварительно изучим свойство относительной энтропии (расхождения Кульбака–Лейблера)

$$D(p\|q) = \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i},$$

где n – произвольное натуральное число, $p = (p_1, \ldots, p_n)$, $q = (q_1, \ldots, q_n)$ – элементы n-мерного симплекса Γ_n (множество всех распределений вероятностей на конечном множестве из n элементов). Предполагаем, что $0 \ln 0 = 0$.

Неравенства $p > q$, $p \geq q$, $p \geq 0$ на векторах будем понимать покомпонентно; здесь 0 – вектор с нулевыми координатами. В дальнейшем будут использоваться следующие свойства относительной энтропии. Для любых $p, q, w \in \Gamma_n$ таких, что $q, w > 0$, будет

$$D(p\|q) \leq D(p\|w) + \ln\left(\sum_{i=1}^{n} p_i \frac{w_i}{q_i}\right).$$

Если $q \geq \mu w$ для некоторого числа $\mu > 0$, то $D(p\|q) \leq D(p\|w) + \ln \frac{1}{\mu}$. В частности, при $p = w$ будет $D(w\|q) \leq \ln \frac{1}{\mu}$ при $q \geq \mu w$. Действительно, из вогнутости логарифма будет

$$D(p\|q) - D(p\|w) = \sum_{i=1}^{n} p_i \ln \frac{w_i}{q_i} \leq \ln\left(\sum_{i=1}^{n} p_i \frac{w_i}{q_i}\right). \tag{2}$$

Если к тому же $q \geq \mu w$, будет $\sum_{i=1}^{n} p_i \frac{w_i}{q_i} \leq \sum_{i=1}^{n} p_i \frac{w_i}{\mu w_i} = \frac{1}{\mu}$.

Отсюда получаем важное для дальнейшего следствие. Пусть $p \in \Gamma_{t+1}$ и $q = \sum_{i=0}^{t} \beta_i w_i$, где $w_i \in \Gamma_{t+1}$, $w_i > 0$ при $0 \leq i \leq t$, $\beta = (\beta_0, \ldots, \beta_t) \in \Gamma_{t+1}$ и $\beta > 0$. Тогда

$$D(p\|q) \leq D(p\|w_i) + \ln \frac{1}{\beta_i}$$

для любого i. В частности, при $p = q$ имеем оценку расхождения между произвольным элементом w_i выпуклой комбинацией и вектором выпуклой комбинации:

$$D\left(w_i \| \sum_{i=0}^{n} \beta_i w_i\right) \leq \ln \frac{1}{\beta_i}$$

для любого i.

В дальнейшем на каждом шаге t вектор $w_t^m = (w_{1,t}^m, \ldots, w_{N,t}^m)$ будет представлять накопленные нормированные веса экспертов на этом шаге. Новое апостериорное распределение $w_{t+1} = (w_{1,t+1}, \ldots, w_{N,t+1})$ экспертов (для использования на шаге $t+1$) будет определяться в виде выпуклой комбинации $w_{t+1} = \sum_{s=0}^{t} \beta_s^{t+1} w_s^m$ с весами β_s^{t+1}, $0 \leq s \leq t$, где $w_s^m = (w_{1,s}^m, \ldots, w_{N,s}^m)$ – вектор нормированных весов экспертов $\{1, \ldots, N\}$ на шаге s. В координатах эта операция записывается в виде $w_{i,t+1} = \sum_{s=0}^{t} \beta_s^{t+1} w_{i,s}^m$ при $1 \leq i \leq N$.

Вектор $\beta^{t+1} = (\beta_0^{t+1}, \ldots, \beta_t^{t+1})$ называется *схемой смешивания*. Для него выполнено $\sum_{s=0}^{t} \beta_s^{t+1} = 1$ и $\beta_s^{t+1} \geq 0$ при $0 \leq s \leq t$. Вектор β^{t+1} задает веса, с помощью которых смешиваются прошлые распределения экспертов, и может заново задаваться на каждом шаге t.

Далее представим алгоритм смешивания апостериорных распределений экспертов МРР. В отличие от ранее рассмотренных методов экспоненциального смешивания этот алгоритм использует не только текущие накопленные веса экспертов, но и смешивает эти веса и все веса, накопленные на прошлых шагах. Это позволяет не забывать прошлые удачные комбинации экспертов, даже если локально на данном шаге эти эксперты несли большие потери. Если в будущем такие эксперты вновь будут нести малые потери, их веса быстро возрастут.

Алгоритм смешивания апостериорных распределений экспертов с адаптивным параметром обучения.

Полагаем $w_{i,1} = w^m_{i,0} = \frac{1}{N}$ при $i = 1, \ldots, N$, $\Delta_0 = 0$, $\eta_1 = \infty$.
FOR $t = 1, \ldots, T$
Предсказываем распределение потерь с весами $w_t = (w_{1,t}, \ldots, w_{N,t})$.
Получаем потери экспертов $l_t = (l^1_t, \ldots, l^N_t)$.
Вычисляем потери алгоритма $h_t = \sum_{i=1}^N w_{i,t} l^i_t$.
Переопределяем веса всех экспертов и параметр обучения в три этапа:
Loss Update
$w^m_{i,t} = \frac{w_{i,t} e^{-\eta_t l^i_t}}{\sum_{j=1}^N w_{j,t} e^{-\eta_t l^j_t}}$ при $1 \le i \le N$.
Mixing Update
Выбираем схему смешивания $\beta^{t+1} = (\beta^{t+1}_0, \ldots, \beta^{t+1}_t)$ и определим
$w_{i,t+1} = \sum_{s=0}^t \beta^{t+1}_s w^m_{i,s}$ при $1 \le i \le N$.
Learning Parameter Update
Вычисляем экспоненциально смешанные потери $m_t = -\frac{1}{\eta_t} \ln \sum_{i=1}^N w_{i,t} e^{-\eta_t l^i_t}$ и величины $\delta_t = h_t - m_t$ и $\Delta_t = \Delta_{t-1} + \delta_t$. После этого полагаем $\eta_{t+1} = \ln(eN)/\Delta_t$ – значение параметра обучения для использования на следующем шаге. Здесь e – основание натурального алгоритма.
ENDFOR

Величины h_t, m_t, δ_t и Δ_t определены в алгоритме. Из выпуклости экспоненты $m_t \le h_t$ для всех t; также $\Delta_t \le \Delta_{t+1}$. Пусть $M_T = \sum_{t=1}^T m_t$ – кумулятивные экспоненциально смешанные потери за первые T шагов. Из определения $H_T = M_T + \Delta_T$.

Будем использовать некоторые схемы смешивания из [10].

Пример 1. Вводим последовательность параметров $1 \ge \alpha_1 \ge \alpha_2 \ge \cdots \ge 0$. В дальнейшем будем использовать $\alpha_t = \frac{1}{t}$. Полагаем $\beta^{t+1}_t = 1 - \alpha_{t+1}$ и $\beta^{t+1}_0 = \alpha_{t+1}$ ($\beta^{t+1}_s = 0$ при $0 < s < t$). В этом случае прогноз алгоритма (веса экспертов) для шага $t+1$ равен

$$w_{i,t+1} = \frac{\alpha_{t+1}}{N} + (1 - \alpha_{t+1}) w^m_{i,t}$$

для всех i и t. Соответствующий алгоритм называется Fixed Share; он был предложен в [9] и изучался в работах [6, 14].

Пример 2. Равномерное влияние прошлого. Полагаем $\beta^{t+1}_t = 1 - \alpha_{t+1}$ и $\beta^{t+1}_s = \frac{\alpha_{t+1}}{t}$, при $0 \le s < t$. В этом случае прогноз алгоритма (веса экспертов) для шага $t+1$ равен

$$w_{i,t+1} = \alpha_{t+1} \sum_{s=0}^{t-1} \frac{w^m_{i,s}}{t} + (1 - \alpha_{t+1}) w^m_{i,t}$$

для всех i и t.

Будем использовать более общее понятие регрета. Ранее сравнивались кумулятивные потери смешивающего алгоритма с кумулятивными потерями произвольного эксперта. Теперь будем сравнивать потери смешивающего алгоритма с потерями заданной выпуклой комбинации потерь экспертов.

Обозначим через $l_t = (l^1_t, \ldots, l^N_t)$ вектор потерь всех экспертов на шаге t. Будем рассматривать выпуклые комбинации потерь экспертов на шаге t: $(q_t \cdot l_t) = \sum_{i=1}^N q_{i,t} l^i_t$, где $q_t = (q_{1,t}, \ldots, q_{N,t})$ – вектор сравнения потерь, $q_t \in \Gamma_N$. В частности, при единичном векторе сравнения $q_t = (0, \ldots, 1, \ldots, 0)$ будет $(q_t \cdot l_t) = l^i_t$, где 1 является i-й координатой q_t.

Следуя [15], рассмотрим еще более общее понятие регрета – разность между кумулятивными потерями алгоритма и суммой выпуклых комбинаций потерь экспертов относительно заданной последовательности $q_1, \ldots, q_T$ векторов сравнений

$$R_T^q = \sum_{t=1}^{T} h_t - \sum_{t=1}^{T} (q_t \cdot l_t).$$

При $q_1 = \cdots = q_T = e_i$, где $e_i = (0, \ldots, 1, \ldots, 0)$ – единичный вектор, получаем регрет $R_T^i = H_T - L_T^i$ относительно эксперта i.

Дальнейшие рассуждения используют лемму из [10].

Лемма 1. *Для произвольных t, $\eta_t > 0$ и произвольного вектора сравнения $q_t \in \Gamma_N$ будет*

$$m_t = (q_t \cdot l_t) + \frac{1}{\eta_t}(D(q_t \| w_t) - D(q_t \| w_t^m)). \tag{3}$$

Доказательство. Используя (2), получим неравенство (3) следующими преобразованиями:

$$\begin{aligned} m_t - \sum_{i=1}^{N} q_{i,t} l_t^i &= \sum_{i=1}^{N} q_{i,t} \left(\frac{1}{\eta_t} \ln e^{-\eta_t l_t^i} + m_t \right) = \\ &= \frac{1}{\eta_t} \sum_{i=1}^{N} q_{i,t} \left(\ln e^{-\eta_t l_t^i} - \ln \sum_{j=1}^{N} w_{j,t} e^{-\eta_t l_t^j} \right) = \\ &= \frac{1}{\eta_t} \sum_{i=1}^{N} q_{i,t} \ln \frac{e^{-\eta_t l_t^i}}{\sum_{j=1}^{N} w_{j,t} e^{-\eta_t l_t^j}} = \\ &= \frac{1}{\eta_t} \sum_{i=1}^{N} q_{i,t} \ln \frac{w_{i,t}^m}{w_{i,t}} = \frac{1}{\eta_t}(D(q_t \| w_t) - D(q_t \| w_t^m)). \end{aligned}$$

$\triangle$

Применим лемму для схемы смешивания из примера 1: $\beta_t^{t+1} = 1 - \alpha_{t+1}$, $\beta_s^{t+1} = 0$ при $0 < s < t$ и $\beta_0^{t+1} = \alpha_{t+1}$. Веса вычисляются по формуле $w_{i,t+1} = \frac{\alpha_{t+1}}{N} + (1 - \alpha_{t+1}) w_{i,t}^m$ для всех i и t. Далее полагаем $\alpha_t = \frac{1}{t}$ для всех $t = 1, 2, \ldots$

Рассматриваем произвольную последовательность $l_t = (l_t^1, \ldots, l_t^N)$ векторов потерь экспертов и произвольную последовательность $q_t = (q_{1,t}, \ldots, q_{N,t})$ векторов сравнения потерь экспертов на всех шагах.

Предположим, что вектор сравнения q_t изменяется k раз при $1 \le t \le T$. Пусть $1 < t_1 < t_2 < \cdots < t_k$ – те шаги, на которых происходят изменения, т.е. $q_{t_j} \ne q_{t_j - 1}$ и $q_t = q_{t-1}$ для всех остальных шагов, где $t > 1$. Полагаем $t_0 = 1$ и $t_{k+1} = T + 1$. Напомним, что $w_{i,1} = w_{i,0}^m = \frac{1}{N}$ при $i = 1, \ldots, N$.

Суммируем неравенство (3) на произвольном интервале постоянства вектора сравнения $q_t = q_{t-1}$ при $t_j + 1 \le t \le t_{j+1} - 1$ и получаем

$$\sum_{t=t_j+1}^{t_{j+1}-1} m_t = \sum_{t=t_j+1}^{t_{j+1}-1} (q_t \cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1} \frac{1}{\eta_t}(D(q_t\|w_t) - D(q_t\|w_t^m)) =$$

$$= \sum_{t=t_j+1}^{t_{j+1}-1} (q_t \cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1} \left(\frac{1}{\eta_{t-1}} D(q_t\|w_t) - \frac{1}{\eta_t} D(q_t\|w_t^m)\right) +$$

$$+ \sum_{t=t_j+1}^{t_{j+1}-1} \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) D(q_t\|w_t) \le \tag{4}$$

$$\le \sum_{t=t_j+1}^{t_{j+1}-1} (q_t \cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1} \left(\frac{1}{\eta_{t-1}} D(q_t\|w_t) - \frac{1}{\eta_t} D(q_t\|w_t^m)\right) +$$

$$+ \sum_{t=t_j+1}^{t_{j+1}-1} \frac{1}{\ln N + 1} \delta_{t-1} \left(\ln N + \ln \frac{1}{\alpha_t}\right) \le \tag{5}$$

$$\le \sum_{t=t_j+1}^{t_{j+1}-1} (q_t \cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1} \left(\frac{1}{\eta_{t-1}} D(q_t\|w_{t-1}^m) - \frac{1}{\eta_t} D(q_t\|w_t^m)\right) +$$

$$+ \sum_{t=t_j+1}^{t_{j+1}-1} \frac{1}{\eta_{t-1}} \ln \frac{1}{1-\alpha_t} + \sum_{t=t_j+1}^{t_{j+1}-1} \left(\frac{1}{\ln N + 1} \ln \frac{1}{\alpha_t} + 1\right) \delta_{t-1} \le \tag{6}$$

$$\le \sum_{t=t_j+1}^{t_{j+1}-1} (q_t \cdot l_t) + \left(\frac{1}{\eta_{t_j}} D(q_{t_j}\|w_{t_j}^m) - \frac{1}{\eta_{t_{j+1}-1}} D(q_{t_j}\|w_{t_{j+1}-1}^m)\right) +$$

$$+ \sum_{t=t_j+1}^{t_{j+1}-1} \frac{1}{\eta_t} \ln \frac{1}{1-\alpha_t} + \sum_{t=1}^{t_{j+1}-1} \delta_{t-1} \ln \frac{1}{\alpha_t} + \sum_{t=t_j+1}^{t_{j+1}-1} \delta_{t-1}. \tag{7}$$

При переходе от (4) к (5) было использовано неравенство $w_{i,t} \ge \frac{\alpha_t}{N}$ для произвольного t. Отсюда

$$D(q_t\|w_t) = \sum_{i=1}^{N} q_{i,t} \ln \frac{q_{i,t}}{w_{i,t}} \le \sum_{i=1}^{N} q_{i,t} \ln q_{i,t} - \sum_{i=1}^{N} q_{i,t} \ln \frac{\alpha_t}{N} \le \ln N + \ln \frac{1}{\alpha_t}. \tag{8}$$

При переходе от (5) к (6) было использовано (2) при $i = t - 1$:

$$D(q_t\|w_t) \le D(q_t\|w_{t-1}^m) + \ln \frac{1}{1-\alpha_t}.$$

При переходе от (6) к (7) было произведено сокращение соседних слагаемых, в результате чего остались первое и последнее слагаемые суммы.

Для шагов t, где происходили изменения вектора сравнения $t = t_1, \ldots, t_k$, полагаем $s = 0$, $\beta_0^{t_j} = \alpha_{t_j}$, т.е.

$$m_{t_j} \le (q_{t_j} \cdot l_{t_j}) + \frac{1}{\eta_{t_j}} D(q_{t_j}\|w_0^m) - \frac{1}{\eta_{t_j}} D(q_{t_j}\|w_{t_j}^m) + \frac{1}{\eta_{t_j}} \ln \frac{1}{\alpha_{t_j}}. \tag{9}$$

Складываем эти неравенства и получаем

$$M_T - \sum_{t=1}^{T}(q_t \cdot l_t) \le \sum_{j=1}^{k}\left(\frac{1}{\eta_{t_j}} D(q_{t_j}\|w_0^m) - \frac{1}{\eta_{t_{j+1}-1}} D(q_{t_j}\|w_{t_{j+1}-1}^m)\right) + \tag{10}$$

$$+\sum_{t=1}^{T}\frac{1}{\eta_t}\ln\frac{1}{1-\alpha_t} + \sum_{t=1}^{T}\delta_{t-1}\ln\frac{1}{\alpha_t} + \sum_{t=1}^{T}\delta_{t-1} + \sum_{j=1}^{k}\Delta_{t_j-1}\ln\frac{1}{\alpha_{t_j}} \le \tag{11}$$

$$\le ((k+2)\ln T + k + 1)\Delta_T.$$

Для перехода от (10) к (11) были использованы монотонность Δ_t, неотрицательность относительной энтропии $D(q\|p) \ge 0$ для всех q, p и равенство $D(q\|w_0^m) = \ln N$, откуда

$$\sum_{j=1}^{k}\frac{1}{\eta_{t_j}} D(q_{t_j}\|w_0^m) \le k\ln N\Delta_T.$$

Кроме этого, огрублялись неравенства после деления на $\ln eN$. При $\alpha_t = \frac{1}{t}$ используем неравенства

$$\sum_{t=1}^{T}\frac{1}{\eta_t}\ln\frac{1}{1-\alpha_t} \le \frac{1}{\eta_T}\ln T \le \Delta_T \ln T.$$

Таким образом, доказано следующее утверждение.

Теорема 1. *Пусть $\alpha_t = \frac{1}{t}$ для всех t и используется схема смешивания из примера 1. Тогда для любого T, для любой последовательности векторов потерь экспертов и любой последовательности векторов сравнения $q_t \in \Gamma_N$ с не более чем k изменениями, получаемых в режиме онлайн, выполнено неравенство*

$$M_T \le \sum_{t=1}^{T}(q_t \cdot l_t) + ((k+2)\ln T + k + 1)\Delta_T.$$

Кроме этого,

$$H_T \le \sum_{t=1}^{T}(q_t \cdot l_t) + \gamma_{k,T}\Delta_T.$$

Обозначим $\gamma_{k,T} = (k+2)(\ln T + 1)$.

Соответствующая оценка для примера 2 получается аналогичным образом. Так как из определения $w_{i,t} \ge \frac{\alpha_t}{Nt}$, то для произвольного t неравенство (8) изменяется на $D(q_t\|w_t) \le \ln N + \ln T + \ln\frac{1}{\alpha_t}$. Также последнее слагаемое неравенства (9) заменяется на $\frac{1}{\eta_{t_j}}\ln\frac{1}{t\alpha_{t_j}}$. В результате получаем $\gamma_{k,T} = (2k+3)\ln T + k + 2$.

Напомним обозначения $l_t^- = \min_{1\le i\le N} l_t^i$, $l_t^+ = \max_{1\le i\le N} l_t^i$, $s_t = l_t^+ - l_t^-$, $L_T^- = \sum_{t=1}^{T} l_t^-$, $L_T^+ = \sum_{t=1}^{T} l_t^+$, $S_T = \max_{1\le t\le T} s_t$ и $L_T^{(k)} = \sum_{t=1}^{T}(q_t \cdot l_t)$, где $q_1, \ldots, q_T$ – некоторая заданная последовательность векторов сравнения и k – число шагов $t \le T$, для которых $q_t \ne q_{t-1}$.

В разделе 3 приводится вывод верхней оценки величины Δ_T:

$$\Delta_T \le \sqrt{S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}\ln(eN)} + \left(\gamma_{k,T}(\ln N + 2) + \frac{2}{3}\ln N + \frac{5}{3}\right)S_T$$

для произвольного T и $1 \le k \le T$.

Используя эту оценку, можно сформулировать основной результат статьи.

Теорема 2. *Для произвольного T и для произвольной последовательности векторов сравнения $q_t \in \Gamma_N$ с не более чем k изменениями, получаемых в режиме онлайн, будет выполнено неравенство*

$$R_T^{(k)} = H_T - L_T^{(k)} \le \gamma_{k,T}\sqrt{S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}(\ln N + 1)} + \\ +\gamma_{k,T}\left(\gamma_{k,T}(\ln N + 2) + \frac{2}{3}\ln N + \frac{5}{3}\right)S_T, \tag{12}$$

где $\gamma_{k,T} = (k+2)(\ln T + 1)$.

Доказательство теоремы 2 приведено в разделе 3.

В частности, такая же оценка имеет место для составных экспертов $i_1, \ldots, i_T$ (т.е. последовательностей исходных экспертов $1 \le i_t \le N$ при $1 \le t \le T$).

Сложность составного эксперта $s(i_1, \ldots, i_T)$ определяется как число шагов $1 < t \le T$, на которых $i_{t-1} \ne i_t$. Пусть также в качестве векторов сравнения q_t используются единичные векторы, причем, $q_{t-1} \ne q_t$ для не более чем k различных $1 < t \le T$. Тогда так как для некоторой последовательности единичных векторов q_t будет

$$L_T^{(k)} = \sum_{t=1}^{T}(q_t \cdot l_t) = \min_{s(i_1,\ldots,i_T)\le k}\sum_{t=1}^{T} l_t^{i_t},$$

то неравенство (12) можно переписать в виде

$$H_T \le \min_{s(i_1,\ldots,i_T)\le k}\sum_{t=1}^{T} l_t^{i_t} + \gamma_{k,T}\sqrt{S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}(\ln N + 1)} + \\ +\gamma_{k,T}\left(\gamma_{k,T}(\ln N + 2) + \frac{2}{3}\ln N + \frac{5}{3}\right)S_T.$$

3. ТЕХНИЧЕСКИЕ ОЦЕНКИ

Доказательство теоремы 2, в целом, аналогично доказательству теоремы 8 из [8], за исключением того, что веса экспертов определяются несколько иначе. Приведем необходимые оценки с помощью неравенства Бернштейна через кумулятивную дисперсию потерь экспертов $V_T = \sum_{t=1}^{T} v_t$, где

$$v_t = Var_{j\sim w_t}[l_t^j] = \mathrm{E}_{j\sim w_t}[(l_t^j - \mathrm{E}_{j\sim w_t}[(l_t^j)])^2] = \sum_{j=1}^{N} w_t(l_t^j - h_t)^2,$$

$w_t = (w_{1,t}, \ldots, w_{N,t})$ – вектор нормированных весов экспертов.

Используем обозначения $\eta_t = \frac{\ln(eN)}{\Delta_{t-1}}$, где $\Delta_t = \sum_{t=1}^{T}\delta_t$ ($\Delta_0 = 0$), $\delta_t = m_t - h_t$, $m_t = -\frac{1}{\eta_t}\ln\sum_{i=1}^{N} w_{i,t}e^{-\eta_t l_t^i}$, $h_t = \sum_{i=1}^{N} w_{i,t}l_t^i$, $w_{i,t}$ – нормированные веса алгоритма MPP. Также $s_t = l_t^+ - l_t^-$, где $l_t^+ = \max_{1\le j\le N} l_t^j$, $l_t^- = \min_{1\le j\le N} l_t^j$.

Лемма 2. *Величина δ_t удовлетворяет неравенству*

$$\delta_t \le \frac{e^{s_t\eta_t} - 1 - s_t\eta_t}{\eta_t s_t^2}v_t. \tag{13}$$

Доказательство. Докажем это выражение, используя неравенство Бернштейна (см. леммы А3–А5 из [6]). Пусть $X \in [0,1]$ – случайная величина и $\sigma = \sqrt{\mathrm{E}X^2 - (\mathrm{E}X)^2}$. Тогда для произвольного $\eta > 0$ будет $\ln \mathrm{E}[e^{-\eta(X-\mathrm{E}X)}] \le \sigma^2(e^\eta - \eta - 1)$. Рассмотрим случайную величину, которая принимает значения l_t^j с вероятностями $w_{j,t}$, где $j = 1, \ldots, N$. Преобразуем ее так, чтобы ее значения принадлежали отрезку $[0,1]$: $X_t^j = \dfrac{l_t^j - l_t^-}{s_t}$. Тогда неравенство Бернштейна можно записать следующим образом:

$$\ln \mathrm{E}_{j\sim w_t}\left(e^{-\eta\left(X_t^j - \mathrm{E}X_t^j\right)}\right) \le \sigma^2\left(e^\eta - 1 - \eta\right) \tag{14}$$

для произвольного $\eta > 0$.

Запишем левую часть неравенства (14) более детально при $\eta = s_t\eta_t$:

$$\ln \mathrm{E}_{j\sim w_t}\left(e^{-\eta\left(X_t^j - \mathrm{E}X_t\right)}\right) = \ln\left(\sum_{j=1}^N w_{t,j} e^{-s_t\eta_t\left(\frac{l_t^j - l_t^-}{s_t} - \sum_{i=1}^N w_{i,t}\frac{l_t^i - l_t^-}{s_t}\right)}\right) =$$

$$= \ln\left(\frac{\sum_{j=1}^N w_{j,t} e^{-\eta_t\left(l_t^j - l_t^-\right)}}{e^{-\eta_t\sum_{j=1}^N w_{j,t}\left(l_t^j - l_t^-\right)}}\right) = \ln\left(\frac{\sum_{j=1}^N w_{j,t} e^{-\eta_t l_t^j}}{e^{-\eta_t\sum_{j=1}^N w_{j,t} l_t^j}}\right) =$$

$$= \ln\sum_{j=1}^N w_{j,t} e^{-\eta_t l_t^j} + \eta_t\sum_{j=1}^N w_{j,t} l_t^j = \eta_t(h_t - m_t) = \eta_t\delta_t.$$

Тогда неравенство (14) записывается в виде

$$\eta_t\delta_t \le \sigma^2\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right) = Var_{j\sim w_t}[X_t^j]\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right) =$$

$$= \frac{1}{s_t^2} Var_{j\sim w_t}[l_t^j]\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right).$$

Отсюда получаем необходимое неравенство $\delta_t \le \dfrac{e^{s_t\eta_t} - 1 - s_t\eta_t}{\eta_t s_t^2} v_t$. $\triangle$

Запишем оценку леммы 2 в виде

$$\delta_t \le \frac{g(s_t\eta_t)}{s_t} v_t, \text{ где } g(x) = \frac{e^x - x - 1}{x}. \tag{15}$$

Воспроизведем оценку на кумулятивную смешанную разность Δ_T из работы [8]. Используем $\eta_t = \frac{\ln(eN)}{\Delta_{t-1}}$.

Лемма 3. *Кумулятивная смешанная разность Δ_T удовлетворяет неравенству $(\Delta_T)^2 \le \ln(eN)V_T + \left(\frac{2}{3}\ln(eN) + 1\right) S_T\Delta_T$, где $S_T = \max_{1\le t\le T}(l_t^+ - l_t^-)$.*

Доказательство. Распишем квадрат кумулятивной смешанной разности:

$$(\Delta_T)^2 = \sum_{t=1}^T\left(\Delta_t^2 - \Delta_{t-1}^2\right) = \sum_{t=1}^T\left(\left(\Delta_{t-1} + \delta_t\right)^2 - \Delta_{t-1}^2\right) =$$

$$= \sum_{t=1}^T\left(2\delta_t\Delta_{t-1} + \delta_t^2\right) = \sum_{t=1}^T\left(\frac{2\delta_t}{\eta_t}\ln(eN) + \delta_t^2\right) \le$$

$$\le \sum_{t=1}^T\left(\frac{2\delta_t}{\eta_t}\ln(eN) + s_t\delta_t\right) \le 2\ln(eN)\sum_{t=1}^T\frac{\delta_t}{\eta_t} + S_T\Delta_T. \tag{16}$$

Получим оценку $\frac{\delta_t}{\eta_t}$ с помощью неравенства (15). Перепишем его в виде

$$\frac{1}{2}v_t \geq \frac{\delta_t s_t}{2g\left(s_t\eta_t\right)} = \frac{\delta_t}{\eta_t} + A,$$

$$A = \frac{\eta_t s_t \delta_t - 2g\left(s_t\eta_t\right)\delta_t}{2g\left(s_t\eta_t\right)\eta_t} = \frac{s_t\left(\eta_t^2\delta_t s_t^2 - 2\delta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)\right)}{2s_t\eta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)} =$$

$$= s_t\delta_t\frac{\frac{1}{2}\left(s_t\eta_t\right)^2 - e^{s_t\eta_t} + s_t\eta_t + 1}{s_t\eta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)} = -\varphi\left(s_t\eta_t\right)s_t\delta_t,$$

где $\varphi(x) = \frac{e^x - \frac{1}{2}x^2 - x - 1}{xe^x - x^2 - x}$.

Нетрудно оценить $\varphi(x) \leq 1/3$.[1]

Откуда

$$\frac{\delta_t}{\eta_t} \leq \frac{1}{3}s_t\delta_t + \frac{1}{2}v_t. \tag{17}$$

Подставим оценку (17) в неравенство (16) и, суммируя, получим:

$$\left(\Delta_T\right)^2 \leq \ln(eN)V_T + \left(\frac{2}{3}\ln(eN) + 1\right)S_T\Delta_T.$$

$\triangle$

Оценим величину V_T. Из определения $v_t \leq (l_t^+ - h_t)(h_t - l_t^-) \leq \frac{s_t^2}{4}$.

Лемма 4. *Пусть $k \leq T$, $S_T = \max_{1\leq t\leq T} s_t$ и $L_T^{(k)} \leq H_T$. Тогда $V_T \leq S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-} + \gamma_{k,T}S_T\Delta_T$.*

Доказательство. Имеют место неравенства

$$V_T = \sum_{t=1}^{T} v_t \leq \sum_{t=1}^{T}(l_t^+ - h_t)(h_t - l_t^-) \leq S_T\sum_{t=1}^{T}\frac{(l_t^+ - h_t)(h_t - l_t^-)}{s_t} =$$

$$= S_T T\sum_{t=1}^{T}\frac{1}{T}\frac{(l_t^+ - h_t)(h_t - l_t^-)}{s_t} \leq S_T\frac{(L_T^+ - H_T)(H_T - L_T^-)}{L_T^+ - L_T^-}.$$

Последнее неравенство получено применением неравенства Йенсена к вогнутой функции $B(x,y,z) = (z-y)(y-x)/(z-x)$ на множестве $x \leq y \leq z$ (см. [8]). $\triangle$

Согласно условию леммы 4 имеет место неравенство $L_T^{(k)} \leq H_T \leq L_T^{(k)} + \gamma_{k,T}\Delta_T$. Отсюда

$$V_T \leq S_T\frac{(L_T^+ - H_T)(H_T - L_T^-)}{L_T^+ - L_T^-} \leq$$

$$\leq S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} + \gamma_{k,T}\Delta_T - L_T^-)}{L_T^+ - L_T^-} \leq$$

$$\leq S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-} + \gamma_{k,T}S_T\Delta_T. \tag{18}$$

[1] Найдем оценку на $\varphi(x)$, используя ее разложение по формуле Тейлора:

$$\frac{e^x - \frac{1}{2}x^2 - x - 1}{xe^x - x^2 - x} \sim \frac{1 + x + \frac{x^2}{2} - \frac{x^2}{2} - x - 1 + \frac{x^3}{6} + o(x)}{\frac{x^3}{2} + o(x)} \sim \frac{\frac{x^3}{6} + o(x)}{\frac{x^3}{2} + o(x)} \sim \frac{1}{3}.$$

Обозначим

$$A_T = \frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}.$$

Из неравенства (18) и оценки леммы 3 следует, что

$$\Delta_T^2 \le (S_T A_T + \gamma_{k,T} S_T \Delta_T)\ln(eN) + \left(\gamma_{k,T} + \frac{2}{3}\ln(eN) + 1\right) S_T \Delta_T =$$
$$= S_T A_T \ln(eN) + \left(\gamma_{k,T}(2 + \ln N) + \frac{2}{3}\ln N + \frac{5}{3}\right) S_T \Delta_T.$$

Имеем неравенство вида $\Delta_T^2 \le a + b\Delta_T$, где $a = S_T A_T \ln(eN)$,
$b = (\gamma_{k,T}(2 + \ln N) + \frac{2}{3}\ln N + \frac{5}{3})S_T$. Разрешим это неравенство относительно Δ_T:

$$\Delta_T \le \frac{1}{2}b + \frac{1}{2}\sqrt{b^2 + 4a} \le \frac{1}{2}b + \frac{1}{2}\left(\sqrt{b^2} + \sqrt{4a}\right) =$$
$$= \sqrt{a} + b = \sqrt{S_T A_T \ln(eN)} + \left(\gamma_{k,T}(\ln N + 2) + \frac{2}{3}\ln N + \frac{5}{3}\right) S_T.$$

Напомним, что $R_T^{(k)} = H_T - L_T^{(k)}$. Если $H_T \le L_T^{(k)}$, то $R_T^{(k)} \le 0$ и неравенство (12) автоматически выполнено. В противном случае

$$R_T^{(k)} = H_T - L_T^{(k)} \le \gamma_{k,T}\Delta_T \le$$
$$\le \gamma_{k,T}\sqrt{S_T A_T \ln(eN)} + \gamma_{k,T}\left(\gamma_{k,T}(\ln N + 2) + \frac{2}{3}\ln N + \frac{5}{3}\right) S_T,$$

где $\gamma_{k,T} = (k+2)(\ln T + 1)$. Неравенство (12) верно. На этом доказательство теоремы 2 завершается.

4. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Для иллюстрации полученных теоретических результатов были проведены численные эксперименты. Рисунки 1–3 иллюстрируют результаты эксперимента на искусственных данных. На рис. 1 приведены локальные пошаговые потери трех экспертов. Как видим, пошаговые потери экспертов могут быть нулевыми, могут возрастать или убывать, а также быть положительными или отрицательными. На каждом временном сегменте, состоящем из 100 точек, потери каждого эксперта носят монотонный характер, при переходе к следующему сегменту характер изменения резко меняется. Для усложнения работы агрегирующих алгоритмов в процессе онлайн адаптации на измерения локальных потерь накладывали равномерный шум с нулевым средним и единичной дисперсией.

На рис. 2 изображены траектории кумулятивных потерь этих экспертов за 500 шагов, а также потери агрегирующих алгоритмов AdaHedge и MPP (схема смешивания примера 1). На рис. 3 приведены текущие веса этих экспертов, полученные при использовании алгоритмов AdaHedge и MPP. На каждом шаге кумулятивные потери обоих алгоритмов не превосходят потери наилучшего на данном шаге эксперта, а потери алгоритма MPP не превосходят потери наилучшей комбинации всех трех экспертов.

Во втором эксперименте исходным материалом для исследования послужили данные о поминутных котировках акций на рынке BATS(US) за 2015 г., доступные для скачивания на сайте www.finam.ru. Всего были загружены поминутные цены (закрытия минутного интервала) для 42 акций. Сначала для каждой из них с помощью специального алгоритма по предыстории

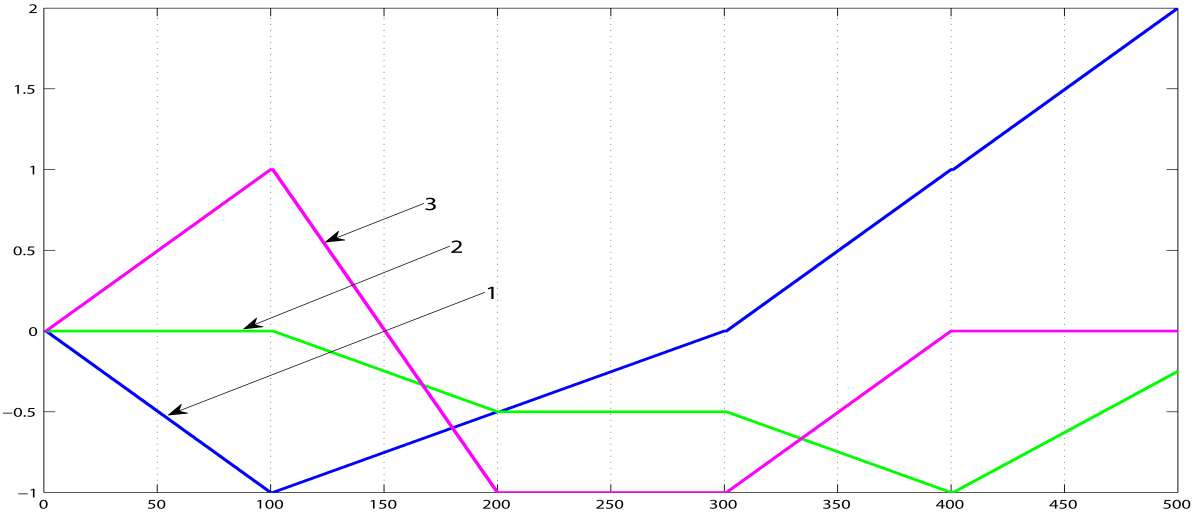

Рис. 1. Пошаговые локальные потери трех искусственных экспертов (1, 2, 3).

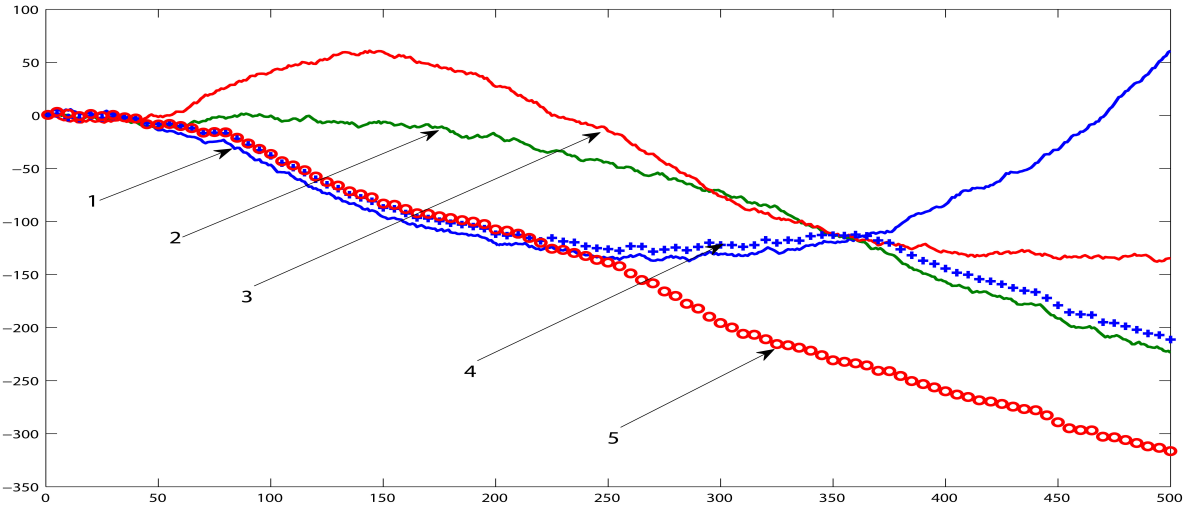

Рис. 2. Кумулятивные потери трех искусственных экспертов (1, 2, 3) в сравнении с потерями агрегирующих алгоритмов AdaHedge (4) и MPP (5).

строили скользящий прогноз цены на начало следующего момента времени. Полученные прогнозы использовали в процедуре виртуальной алгоритмической торговли. Детальное описание работы торгового автомата приведено в [12] и [13]. Коротко суть торгового алгоритма состоит в том, что на каждом шаге игры на основании текущей цены акции и ее прогноза на следующий такт оценивали достижимый доход от входа в рынок в длинную или короткую позицию. Если ожидаемый доход превышал заранее заданный порог, то автомат входил в рынок, а затем, основываясь на следующих изменениях и прогнозах, выбирал момент выхода из рынка. Игра на каждой акции рассматривалась как независимый эксперт. На рис. 4 приведены графики кумулятивных доходов 42 независимых экспертов на каждый рабочий день 2015 г. в сравнении с доходами алгоритмов AdaHedge и MPP. Для удобства сравнения графики приведены в долях от начальной вложенной суммы. На рис. 5 и 6 приведена динамика весов алгоритмов AdaHedge и MPP соответственно.

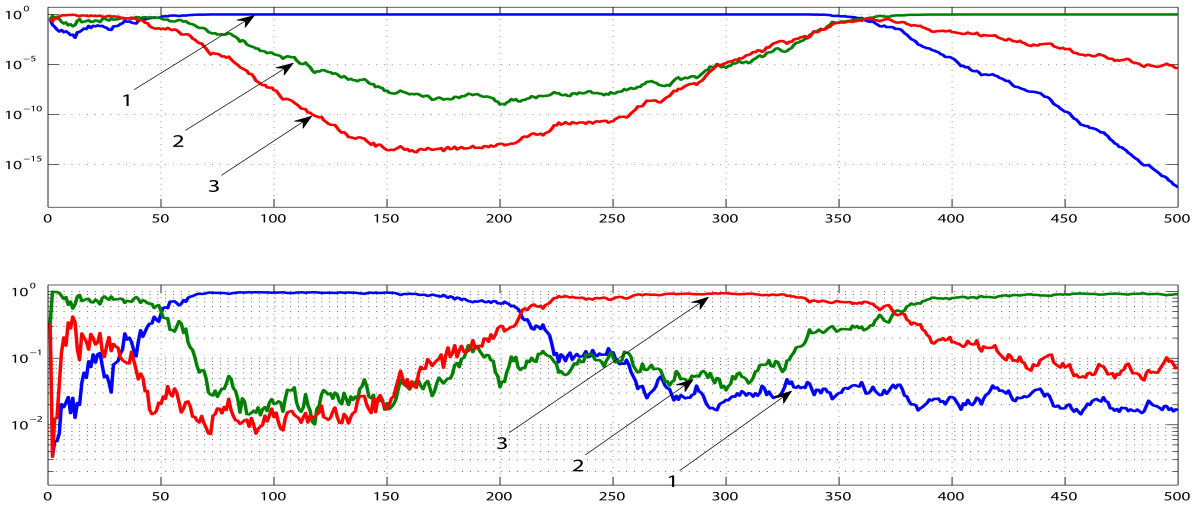

Рис. 3. Динамика изменения весов экспертов (в логарифмической шкале) для алгоритмов AdaHedge (верхний рисунок) и MPP (нижний рисунок).

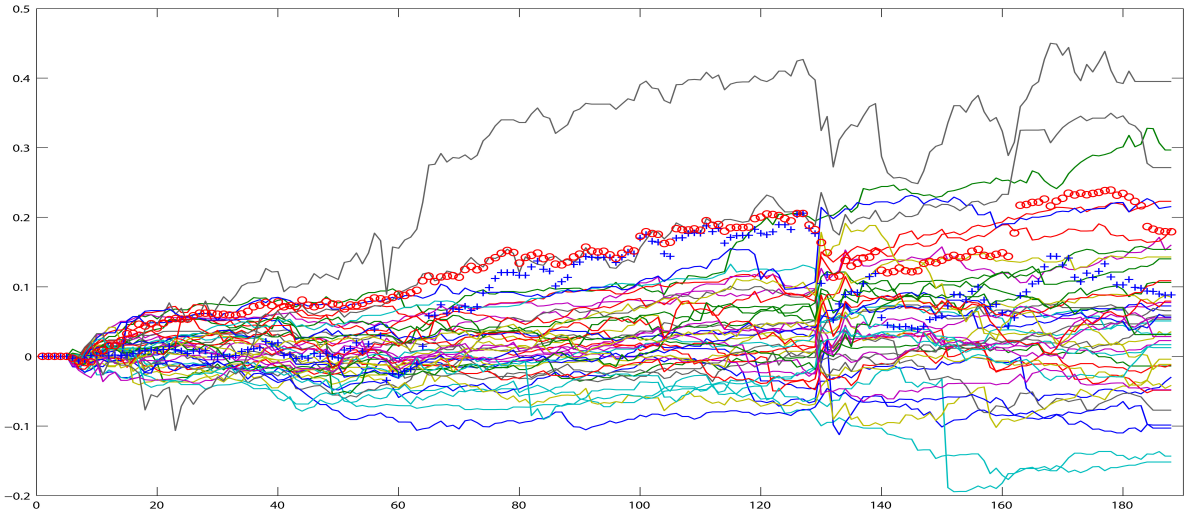

Рис. 4. Кумулятивные доходы 42-х экспертов (тонкие линии) в сравнении с доходами алгоритмов AdaHedge (+++++) и MPP (ooooo) за все рабочие дни 2015 г.

Теоретическая часть статьи основана на сравнении потерь агрегирующего алгоритма с потерями наилучших экспертов. В данном эксперименте максимизировали выигрыш (доход), полученный в процессе игры. Поэтому при использовании алгоритмов смешивания в качестве аргументов подставляли значения доходов с отрицательным знаком.

Агрегирующие алгоритмы AdaHedge и MPP использовали в качестве входных данных только динамические ряды кумулятивных доходов 42 экспертов. Эти алгоритмы в начале каждого дня перераспределяли одну и ту же сумму между 42 экспертами в соответствии с их накопленными весами и в конце дня получали от каждого эксперта соответствующую долю его дохода или убытка. После этого вновь вкладывали ту же фиксированную сумму в экспертные стратегии. Доходы сверх положенной суммы откладывались, а в случае необходимости разрешалось брать в долг. Плата за трансакции не взималась.

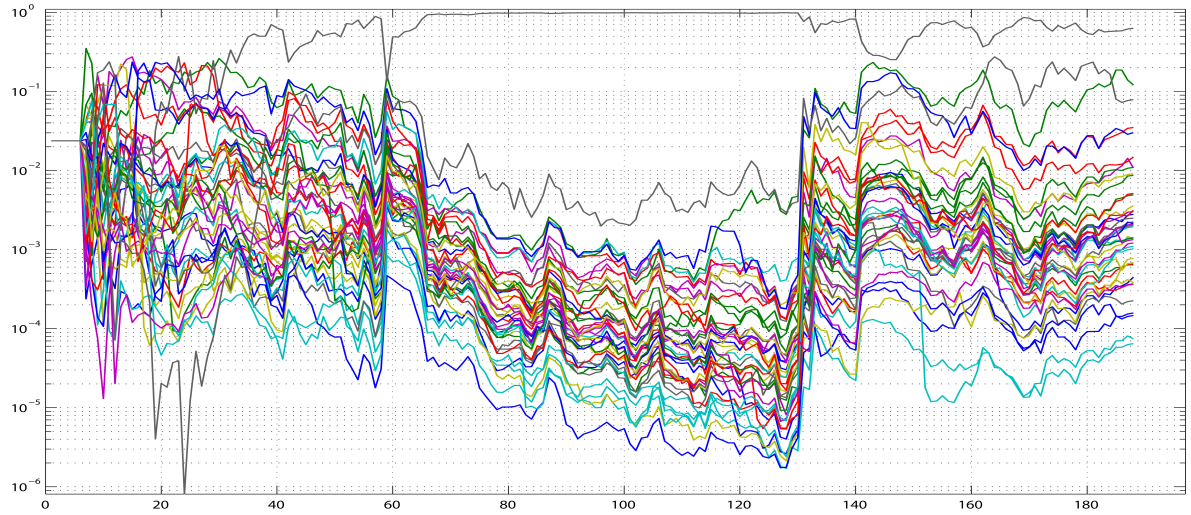

Рис. 5. Динамика изменения весов экспертов (в логарифмической шкале) для алгоритма AdaHedge.

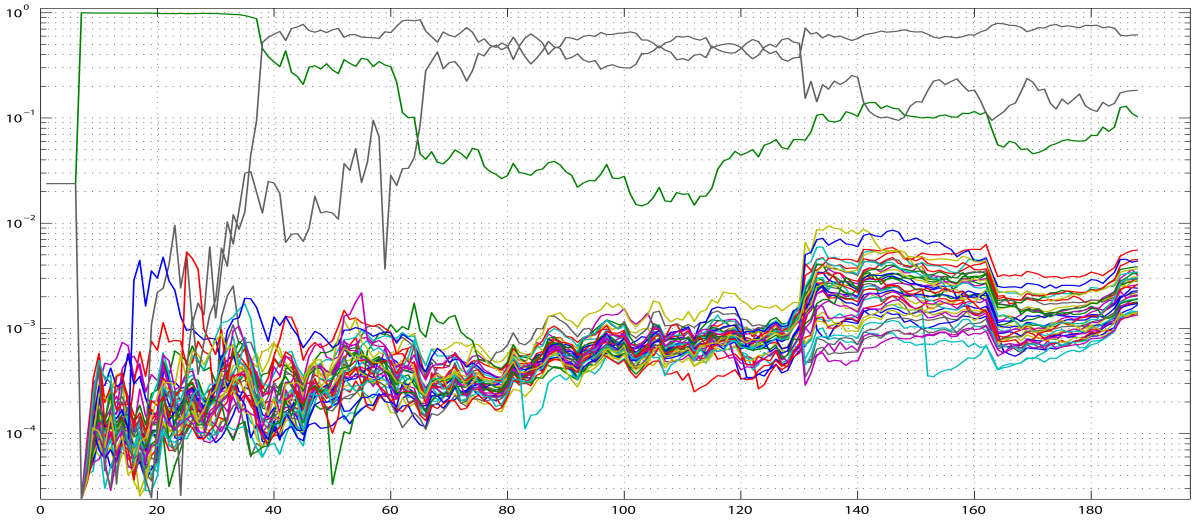

Рис. 6. Динамика изменения весов экспертов (в логарифмической шкале) для алгоритма MPP.

Эксперимент проведен на данных с высокой волатильностью и точками выбросов. Оба эти фактора снижают эффективность обоих алгоритмов по сравнению с искусственными данными.

5. ЗАКЛЮЧЕНИЕ

В данной статье, в рамках теории предсказания (принятия решений) с экспертами в режиме онлайн, предложен адаптивный алгоритм, который агрегирует решения экспертных стратегий и несет потери, не превосходящие (с точностью до некоторой величины, называемой регретом) потери наилучшей комбинации экспертов, произвольным образом распределенных по интервалу прогнозирования. Для построения алгоритма использовался метод адаптивной модификации весов экспертов MPP, который применялся к базовому алгоритму Hedge.

Основным параметром алгоритма Hedge является параметр обучения. В работе использовалась современная версия алгоритма Hedge – алгоритм AdaHedge, в котором параметр обучения вычисляется адаптивно в процессе работы алгоритма (в режиме онлайн). Еще одним преиму-

ществом алгоритма AdaHedge является то, что он не требует заранее задавать границы для потерь экспертных стратегий.

Метод MPP обычно формулируется для постоянного параметра обучения. В данной статье метод MPP сформулирован для переменного параметра обучения, что позволило рассматривать его в комбинации с алгоритмом AdaHedge и получить оценку регрета комбинированного алгоритма.

Предложенный в работе алгоритм обладает преимуществами обоих алгоритмов MPP и AdaHedge. Основной параметр алгоритма – параметр обучения – вычисляется адаптивно в процессе работы алгоритма (в режиме онлайн). Эта особенность определяет численную устойчивость алгоритма в условиях изменяющегося лидерства экспертных стратегий.

Еще одним преимуществом алгоритма является то, что он не требует задания ограничений на потери экспертных стратегий и оценка его регрета проведена в отсутствие этих ограничений. Заметим также, что в рамках данного подхода не делается никаких стохастических предположений об источнике исходных данных. Все оценки регрета даются в терминах минимаксного типа.

Результаты численных экспериментов по смешиванию экспертных решений с помощью предложенных алгоритмов, проведенных на искусственных и реальных данных, показали достаточно высокую степень адаптации алгоритмов AdaHedge и нашего алгоритма в условиях, когда лидерство экспертов часто меняется. При этом наша версия алгоритма MPP быстрее адаптируется к изменяющимся условиям.

Второй эксперимент показывает, что высокая волатильность и резкие скачки в реальных данных снижают эффективность обоих алгоритмов по сравнению с искусственными данными, в которых одношаговые потери экспертов в основном изменяются линейным образом.

СПИСОК ЛИТЕРАТУРЫ

1. *V'yugin V.* Online Aggregation of Unbounded Signed Losses Using Shifting Experts // Proc. Mach. Learn. Res. 2017. V. 60. P. 3–17. http://proceedings.mlr.press/v60/v-yugin17a.html

2. *Littlestone N., Warmuth M.* The Weighted Majority Algorithm // Inform. Comput. 1994. V. 108. P. 212–261.

3. *Freund Y., Schapire R.E.* A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting // J. Comput. Syst. Sci. 1997. V. 55. P. 119–139.

4. *Vovk V.* Aggregating Strategies // Proc. 3rd Annual Workshop on Computational Learning Theory. Fulk M., Case J., eds. San Mateo, CA, Morgan Kaufmann, 1990. P. 371–383.

5. *Vovk V.* A Game of Prediction with Expert Advice // J. Comput. Syst. Sci. Elsevier Publ., 1998. V. 56. No. 2. P. 153–173.

6. *Cesa-Bianchi N., Lugosi G.* Prediction, Learning, and Games. Cambridge Univer. Press, 2006.

7. *Вьюгин В.В.* Математические основы машинного обучения и прогнозирования. М.: Изд.-во МЦНМО, 2018.

8. *de Rooij S., van Erven T., Grunwald P., Koolen W.* Follow the Leader If You Can, Hedge If You Must // J. Mach. Learn. Res. 2004. V. 15. P. 1281–1316.

9. *Herbster M., Warmuth M.* Tracking the best expert // Machine Learning. Springer, 1998. V. 32. No. 2. P. 151–178.

10. *Bousquet O., Warmuth M.* Tracking a Small Set of Experts by Mixing Past Posteriors // J. Mach. Learn. Res. 2002. No. 3. P. 363–396.

11. *Вьюгин В.В., Стельмах И.А., Трунов В.Г.* Адаптивный алгоритм отслеживания наилучшей траектории экспертных решений // Информационные процессы. 2016. Т. 16. № 3. С. 260–280 (*V. V. V'yugin, I. A. Stel'makh, and V. G. Trunov.* Adaptive Algorithm of Tracking the Best Experts Trajectory // Journal of Communications Technology and Electronics, 2017, Vol. 62, No. 12, P.1434–1447. Pleiades Publishing, Inc., 2017).

12. *Вьюгин В.В., Трунов В.Г.* Построение комбинированных финансовых стратегий на основе универсального адаптивного прогнозирования // АиТ. 2016. № 8. С. 136–158. (*V'yugin V. V., Trunov V. G.* Applications of Combined Financial Strategies Based on Universal Adaptive Forecasting // Autom. Remote Control. 2016. V. 77. No. 8. P. 1428–1446).

13. *V'yugin V.* Universal Algorithm for Trading in Stock Market Based on the Method of Calibration // Algorithmic Learning Theory. ALT 2013. Lect. Notes in Computer Science. Jain S., Munos R., Stephan F., Zeugmann T. eds. Berlin, Heidelberg: Springer, 2013. V. 8139. P. 53–67.

14. *Vovk V.* Derandomizing Stochastic Prediction Strategies // Machine Learning. 1999. V. 35. No. 3. P. 247–282. Springer.

15. *Kivinen J., Warmuth M.K.* Averaging Expert Prediction // Proc. Computational Learning Theory: 4th Eur. Conf. (EuroColt '99). Fisher P., Simon H.U. eds. 1999. P. 153–167. Springer.

V.V. V'yugin, V.G. Trunov

Tracking the Best Sequence of Prediction Strategies

In the framework of the decision theoretic online learning, an adaptive algorithm is proposed that aggregates decisions of expert strategies and carries losses that do not exceed (up to a certain value, called regret) losses of the best combination of experts arbitrarily distributed along the prediction time interval. To construct the algorithm, a modified versions of the method of Mixing Past Posteriors and of the algorithm of exponential weighting of expert decisions AdaHedge with adaptive learning parameter was used. An upper bound of the regret of proposed algorithm is obtained. With this approach, no assumptions are used about the nature of data source and experts losses limits. The results of numerical experiments on mixing expert decisions with the help of the proposed algorithm are presented.

KEYWORDS: decision theoretic online learning; prediction with expert advice; mixing past posteriors; adaptive learning rate.