

Использование энтропии в анализе временных рядов (обзор)

А.М. Адельянов,^{1,*} Е.А. Генералов,¹ Вэнь Чжэнь,¹ Л.В. Яковенко¹¹Московский государственный университет имени М.В. Ломоносова, физический факультет
Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2

(Поступила в редакцию 13.03.2024; после доработки 20.05.2024; подписана в печать 29.05.2024)

Результаты экспериментов в области биофизики часто представлены в виде временных рядов, полученных с небольшим разрешением и не всегда большой длины. В частности, при исследованиях воздействий различных физико-химических факторов на бислойные липидные мембраны обычно измеряются трансмембранные ионные токи и их флуктуации. При этом средние значения и дисперсии токов могут значимо не различаться, и по ним определить характер и степень воздействия затруднительно, поэтому разработка подходов к анализу временных рядов никогда не прекращалась. Попытки использовать энтропию распределений случайных величин при таком анализе предпринимались давно, но в практической работе эти подходы были трудно реализуемы, особенно из-за требований к длине рядов и отсутствию шумов. В последние десятилетия в этой области произошли существенные изменения и предложено множество новых методов анализа временных рядов с использованием различных модификаций энтропии. В связи с этим появилась потребность в некой сводке методов, основанных на расчете энтропии, с указанием их достоинств и недостатков. Этой цели и служит предлагаемый краткий обзор энтропийных методов анализа скалярных временных рядов, который может быть полезен при анализе экспериментальных данных. В обзоре рассмотрены только некоторые из базовых подходов, на которых основаны дальнейшие усовершенствования алгоритмов расчетов. Понятие энтропии иногда вызывает затруднения у студентов, поэтому обзор может быть полезен и для использования в педагогических целях.

PACS: 02.50.-r; 02.50.Fz; 01.50.-i УДК: 51.73; 53.05; 53.088; 57.015

Ключевые слова: временные ряды, энтропия, сложность.

DOI: [10.55959/MSU0579-9392.79.2440701](https://doi.org/10.55959/MSU0579-9392.79.2440701)

ВВЕДЕНИЕ

Современные научные исследования требуют новых подходов, в том числе к анализу получаемых данных. По мере развития вычислительных мощностей растет и потребность в методах анализа различных комплексных параметров биологических систем, особенно это актуально для биомедицинских исследований с большими наборами не очень точных данных: анализ передачи сигналов через внутриклеточные сигнальные каскады [1, 2], влияние различных молекул на состояние биологических систем, например организмы [3, 4], а также при дифференцировании нормы от патологии, например при онкологических, кардиологических и нейродегенеративных заболеваниях [5, 6].

Результаты исследований систем различной природы часто представляются в виде временных рядов (ВР) — записей значений какой-либо наблюдаемой величины (сигнала) в зависимости от времени. Такие записи могут быть как дискретными, так и непрерывными, но в последнем случае обычно используется дискретизация с каким-либо постоянным интервалом времени, поэтому далее будут рассмотрены только дискретные ВР. Стационарный

ВР можно рассматривать как отображение некоего состояния системы, его породившей, поэтому исследование ВР в принципе позволяет получить информацию об этой системе. Сам по себе ВР ничего не говорит об устройстве системы, но если существует модель системы с неизвестными параметрами, то анализ ВР позволяет определить по крайней мере число независимых параметров.

Скалярный временной ряд $\{x_i\}_{i=1}^N$ представляет собой массив из N значений наблюдаемой величины $x(t)$, измеренных с постоянным шагом по времени Δt в моменты времени $t_i = t_0 + (i - 1) \Delta t$, где t_0 — момент времени первого измерения, $i = 1, 2, \dots, N$. Обычно анализ ВР связан с решением задач двух типов: идентификация системы, породившей сигнал, и прогноз ее поведения на некоторое время вперед. Под идентификацией понимают оценку каких-либо параметров системы, таких как корреляционная размерность, энтропия, сложность и др. Прогноз имеет целью предсказание будущих значений измеряемого сигнала на основании имеющихся данных, то есть построение аппроксимирующей функции, с помощью которой можно оценить следующее значение измеряемой величины по нескольким предыдущим. Задача прогноза в настоящей статье не рассматривается, основное внимание будет уделено методам оценки стохастичности и сложности ВР.

* E-mail: tashxat@gmail.com

При экспериментальных исследованиях нелинейных химических, гидродинамических и прочих систем часто наблюдается непериодическое изменение их параметров, называемое детерминированным хаосом. На глаз его невозможно отличить от чисто стохастического процесса, поэтому были разработаны математические методы, позволяющие провести классификацию поведения таких систем, которая основана на построении соответствующих фазового портрета. Однако экспериментатору обычно не известны дифференциальные уравнения, представляющие собой модель системы, поэтому возникает проблема построения фазового портрета по ограниченному экспериментальным данным. Эта проблема давно решена, но представляется целесообразным в настоящем обзоре дать очень краткое описание использованных при этом подходов.

Состояние системы характеризуется одной точкой в многомерном (n -мерном) фазовом пространстве, которая со временем описывает некую траекторию. Под фазовым портретом понимают множество таких траекторий, полученных для всех возможных начальных условий при фиксированном наборе управляющих параметров. Сложную задачу построения многомерного фазового портрета можно свести к анализу всего одной наблюдаемой переменной с помощью методов, описанных в [7–9]. Суть их сводится к тому, что почти для любой наблюдаемой величины $B(t)$ и временной задержки τ m -мерный фазовый портрет, построенный с помощью векторов $\mathbf{X}_m(k) = \{B(t_k), B(t_k + \tau), \dots, B(t_k + (m-1)\tau)\}$, где $t_k = k\Delta t$, $k = 1, 2, \dots, \infty$, будет иметь такие же свойства (например, спектр показателей Ляпунова), как и построенный по измерениям n независимых переменных, если $m \geq 2n + 1$. Такой фазовый портрет является вложением исходного многообразия. Задержку τ можно выбирать почти произвольно: например, ее совпадение с периодом в периодическом режиме приведет к вырождению портрета до прямой линии. На практике величину m можно выбрать малой, а затем увеличивать на единицу до тех пор, пока не исчезнут видимые изменения портрета. После построения фазового портрета можно его анализировать разными методами, например построить сечение Пуанкаре, определить максимальный показатель Ляпунова [9], но, поскольку обзор посвящен энтропийным методам, другие подходы здесь рассматриваться не будут.

Энтропийные методы связаны с определением количества и качества информации, получаемой при анализе ВР. Методы определения этих характеристик информации рассматриваются в теории информации, математические основы которой были заложены в работах [10, 11]. Поскольку в них количество информации, необходимое для полного описания исследуемой системы, было отождествлено с уменьшением энтропии распределения вероятных исходов исследования, то в настоящей статье терми-

ны «энтропия» и «информация» будут использоваться взаимозаменяемо, в зависимости от контекста.

Энтропия (информационная) распределения вероятностей и энтропия классической термодинамики — это разные понятия. Основное различие между ними состоит в том, используются ли термодинамические характеристики системы (количество тепла, температура и т.д.) или же характеристики распределений случайных величин, относящихся к системе. Во втором случае энтропия — чисто математическое понятие (и не всегда имеет физический смысл), и ее значение относительно, то есть зависит от того, какие именно случайные величины, относящиеся к системе, исследуются. Так, например, энтропия распределения номиналов монет не зависит от других характеристик монет (год выпуска, дефекты и т.п.) [12]. Связь информационной энтропии с термодинамической последовательно рассмотрена в работах [13, 14].

Суть энтропийных подходов к анализу ВР состоит в том, чтобы, определив тем или иным способом энтропию выборки из ВР, получить информацию о свойствах породившей ВР физической системы. Энтропия является функцией распределения вероятностей случайной величины и изменяется при преобразованиях распределений, поэтому, в принципе, можно подобрать такое преобразование, при котором изменения энтропии наиболее чувствительны к исследуемым характеристикам ВР. Энтропии, соответствующие разным способам расчета, получили разные названия: приближенная энтропия, выборочная энтропия, перестановочная энтропия, дисперсионная энтропия и т.д. Некоторые из них в применении к моновариантным ВР кратко рассмотрены ниже, при этом основное внимание уделено перестановочной энтропии и ее модификациям, которые часто используются при анализе ВР различной природы.

1. ЭНТРОПИЯ ШЕННОНА

В 1948 г. К. Шеннон [10, 11] предложил количественную меру информации, основанную на оценке распределения вероятностей $\{p_i\}_{i=1}^N$ возможных ответов (сообщений) на некий хорошо определенный вопрос Q . В связи с этим важную роль играет нумерация всех возможных ответов и оценка их вероятностей, основанная на имеющихся знаниях X относительно возможных ответов (невозможно оценить количество информации при получении ответа, совершенно не относящегося к вопросу). Хорошо определенный вопрос означает, что вся неопределенность ситуации должна содержаться только в том, какое сообщение будет получено. Функция распределения вероятностей $H(Q/X) = H(p_1, \dots, p_N)$, которая может служить количественной мерой информации и которую Шеннон назвал энтропией, определяется единственным образом, если на нее наложить четыре ограничения (постулаты Шеннона). Количество ин-

формации в сообщении определяется как разность энтропий за счет изменения знаний о возможных ответах до получения сообщения (X) и после получения (X'): $I = H(Q/X) - H(Q/X')$. Таким образом, энтропия является количественной мерой неопределенности знаний о системе; это информация, которой не хватает до полного знания. Энтропия Шеннона лежит в основе всех рассмотренных ниже методов.

Работа Шеннона преследовала в основном практические цели. Математически строгое изложение и обобщение идей Шеннона на случай конечных вероятностных схем впервые было сделано А.Я. Хинчиным [15, 16], поэтому в теории информации принято использовать формулировки основных понятий и теорем теории информации в его интерпретации. Если полная система событий A_1, A_2, \dots, A_N задана вместе с их вероятностями p_1, p_2, \dots, p_N ($p_i \geq 0, \sum_{i=1}^N p_i = 1$), то говорят, что задана конечная схема

$$A = \left(\begin{array}{c} A_1 A_2 \dots A_N \\ p_1 p_2 \dots p_N \end{array} \right).$$

Со всякой конечной схемой связана неопределенность относительно того, какое событие произойдет, поскольку известны лишь вероятности их исходов. Удобной мерой неопределенности конечной схемы является энтропия Шеннона, определяемая уравнением

$$H(p_1, \dots, p_N) = -K \sum_{i=1}^N p_i \ln p_i = - \sum_{i=1}^N p_i \log_a p_i,$$

$$K > 0 \left(\log_a p_i = \frac{\log_b p_i}{\log_b a} = K \log_b p_i \right),$$

где во втором равенстве основание логарифма a может быть любым, но одним и тем же во всех слагаемых (постоянная K связана с выбором этого основания). Энтропия пропорциональна количеству информации, получаемому при реализации конечной схемы A , а от выбора коэффициента пропорциональности зависит единица измерения количества информации. Так, при $K = 1$ и $a = 2$ количество информации измеряется в битах.

Эта функция распределения вероятностей определена единственным образом, если выполняются упомянутые выше постулаты Шеннона–Хинчина:

1. $H(p_1, \dots, p_N)$ непрерывна по всем аргументам;
2. для равномерного распределения $p_i = \frac{1}{N} \forall i$ функция $H(p_1, \dots, p_N) = H(N)$ монотонно возрастает с увеличением N и максимальна по сравнению с другими распределениями вероятностей; если в конечную схему A (с произвольными p_i) добавлено невозможное событие A_{N+1} ($p_{N+1} = 0$), то $H(p_1, \dots, p_{N+1}) = H(p_1, \dots, p_N)$;

3. если заданы две взаимно независимые конечные схемы A и B с числами событий n и m , то совокупность событий $A_k B_l$ ($1 \leq k \leq N, 1 \leq l \leq M$) называется объединением схем A и B и образует новую схему AB с вероятностями событий $p_{kl} = p_k p_l$, при этом $H(AB) = H(A) + H(B)$;

4. для зависимых конечных схем A и B вероятность события B_l в схеме B при условии, что в схеме A произошло событие A_k , будет условной вероятностью $q_{kl} = p(B_l/A_k)$, поэтому $p_{kl} = p_k q_{kl}$; тогда $H(AB) = H(A) + \sum_{k=1}^N p_k \sum_{l=1}^M q_{kl} \log q_{kl} = H(A) + \sum_{k=1}^N p_k H_k(B) = H(A) + H_A(B)$.

В п. 4 последний член в правой части представляет собой математическое ожидание $H(B)$ в схеме A , то есть это математическое ожидание дополнительной информации при реализации схемы B после получения информации при реализации схемы A , поэтому всегда $H(B) \geq H_A(B)$.

Энтропия определена только для дискретных случайных величин. При переходе к непрерывным случайным величинам абсолютное значение энтропии, определенное по аналогии с формулой Шеннона с помощью замены суммирования интегрированием и вероятностей на плотность вероятности, оказывается бесконечным. Это видно, например, из второго постулата Шеннона–Хинчина, так как энтропия неограниченно возрастает с увеличением N , которое происходит при предельном переходе от суммирования к интегрированию. Кроме того, возникает неопределенность из-за того, что под знаком логарифма оказывается размерная величина. Однако в этом случае можно ввести так называемую относительную, или дифференциальную энтропию, которая имеет отличные от энтропии Шеннона свойства и в настоящем обзоре рассматриваться не будет.

Если две исходно не взаимодействовавшие системы A и B с числами микросостояний N_A и N_B объединяются в одну общую систему с числом состояний N_{A+B} , то энтропия такой системы в соответствии с третьим постулатом Шеннона–Хинчина равна сумме энтропий образовавших ее подсистем: $S(AB) = S(A) + S(B)$, то есть энтропия пропорциональна числу элементов системы, когда оно велико. Это выполняется, когда подсистемы статистически независимы (или почти независимы), иначе говоря, когда в системах могут существовать только локальные корреляции. (Пространственные и временные корреляции между микроскопическими параметрами состояния возникают за счет взаимодействий подсистем. Если взаимодействия короткодействующие и такие, что время корреляции и длина корреляции очень малы, то они не скажутся на наблюдаемых макроскопических параметрах. Этот случай здесь назван нами локальными корреляциями.) Аддитивность энтропии означает то же самое, что экстенсивность при условии, что $N_{A+B} = N_A N_B$.

Для случайного ВР любая перестановка его членов не приведет к изменениям моментов распределения, имеющих вид $\sum_i (x_i - \langle x \rangle)^n$, где $\langle x \rangle$ — среднее значение ряда, n — порядок момента, поэтому статистика не дает сведений об организации ВР.

При анализе случайности ВР важно не просто установить, случаен ряд или нет, но оценить степень случайности. Случайность связана со сложностью: простой ВР имеет очевидную структуру, позволяющую точно предсказать следующие его значения, а при максимальной случайности будущие значения непредсказуемы, то есть у стохастического ряда нет структуры. С этим связана возможность «сжатия» (архивирования) исходной информации, содержащейся в ВР: для максимально случайного ВР сжатие невозможно.

Для расчета энтропии нужно оценить вероятности событий p_i , которые на практике обычно определяют по частотам различных исходов, причем в ВР не очень большой длины. Однако это дает смещенную оценку величины энтропии.

1.1. Систематическая ошибка при оценке энтропии по конечной выборке

В биофизических исследованиях очень часто приходится иметь дело со стационарными точечными пуассоновскими процессами (все события независимы друг от друга), поэтому ниже рассмотрена оценка энтропии для такого случая, хотя аналогичная систематическая ошибка возникает и при других распределениях вероятностей случайной величины при частотной оценке вероятностей реализаций [17].

«Эксперимент» состоит в том, что проводятся N измерений случайной величины, значения которой распределены по K интервалам с вероятностями $\mathbf{p} = \{p_1, p_2, \dots, p_K\}$, тогда в i -м интервале окажутся n_i результатов измерений ($i = 1, 2, \dots, K$) и их среднее значение будет равно $\langle n_i \rangle = Np_i$. В соответствии со свойствами распределения Пуассона дисперсия равна среднему значению случайной величины: $\langle (\delta n_i)^2 \rangle = \langle n_i \rangle = Np_i$. Определяя частоту события, как обычно, $f_i = n_i/N$, получим $\langle f_i \rangle = p_i$, $\langle (\delta f_i)^2 \rangle = p_i/N$. Тогда «наивная» оценка энтропии имеет вид: $\tilde{H} = -\sum_{i=1}^K f_i \ln f_i$. Будем измерять энтропию в битах. Разложение \tilde{H} в ряд Тэйлора:

$$\begin{aligned} \tilde{H} &= -\sum_{i=1}^K f_i \log_2 f_i = -\sum_{i=1}^K (p_i + \delta f_i) \log_2 (p_i + \delta f_i) \approx \\ &\approx -\sum_{i=1}^K p_i \log_2 p_i - \sum_{i=1}^K \left(\log_2 p_i + \frac{1}{\ln 2} \right) \delta f_i - \\ &\quad - \frac{1}{2} \sum_{i=1}^K \left(\frac{1}{p_i \ln 2} \right) (\delta f_i)^2 + \dots \end{aligned}$$

После усреднения второй член в правой части об-

нуляется, а третий дает смещение оценки:

$$\begin{aligned} \langle \tilde{H} \rangle &= H - \frac{1}{2 \ln 2} \sum_{i=1}^K \left(\frac{\langle (\delta f_i)^2 \rangle}{p_i} \right) = \\ &= H - \frac{1}{2 \ln 2} \sum_{i=1}^K \left(\frac{p_i}{N p_i} \right) = H - \frac{K}{(2 \ln 2) N}. \end{aligned}$$

Таким образом, при «наивном» подсчете оценки энтропии всегда меньше точного значения и систематическая ошибка пропорциональна числу возможных значений случайной величины и обратно пропорциональна количеству испытаний [17].

2. СЛОЖНОСТЬ, ЭНТРОПИЯ КОЛМОГорова

Существуют два крайних типа систем: стохастические и детерминированные. Эволюция стохастической системы сопровождается ее случайными переходами между микросостояниями, при этом предсказать, в какое состояния система перейдет, невозможно. Поведение стохастической системы невоспроизводимо. Для неустойчивых сложных нелинейных динамических систем при размерности фазового пространства не менее трех возможно появление странного аттрактора и переход в режим динамического хаоса. Внешне поведение хаотической системы не отличается от стохастического, но оно воспроизводимо.

Для описания сложности хаотического поведения динамической системы Колмогоров, Соломонов и Синай предложили использовать K -энтропию, которую определяют следующим образом [18].

F -мерное фазовое пространство системы разбивается на ячейки ε^F с номерами (i_1, \dots, i_d) . Пусть есть аттрактор с траекторией $\mathbf{x}(t)$. Состояния системы измеряют с интервалом τ ; если совместная вероятность того, что $\mathbf{x}(t = \tau) \in i_1, \dots, \mathbf{x}(t = d\tau) \in i_d$ равна $p(i_1, \dots, i_d)$, то:

$$\begin{aligned} K &= -\lim_{\tau \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \lim_{d \rightarrow \infty} \frac{1}{d\tau} \times \\ &\quad \times \sum_{i_1, \dots, i_d} p(i_1, \dots, i_d) \ln p(i_1, \dots, i_d). \end{aligned}$$

$K = 0$, $K \rightarrow \infty$ и $K \neq 0$ в упорядоченной, стохастической и хаотической системах соответственно. K -энтропия — мера скорости потери информации. Практическое использование K -энтропии затруднительно из-за предельных переходов и необходимости очень больших объемов данных.

Количественная оценка сложности ВР связана с решением некоторых проблем теории динамических систем, статистики и вычислительной математики, таких как оценка сложности турбулентного

течения и ее эволюции, выбор более простой модели, описывающей эксперимент, оценка сложности алгоритма. В связи с этим в разных областях исследований возникли разные определения сложности. Наиболее часто используется представление об алгоритмической сложности (по Колмогорову), в которой сложность ВР связана с минимальной длиной бинарной компьютерной программы, которая может его воспроизвести, деленной на размер ВР (вообще говоря, в пределе бесконечного ВР). Однако нет алгоритма, который бы позволил оценить таким образом сложность всех наборов данных, поэтому были предложены другие способы определения сложности.

Алгоритмическая сложность тесно связана с мерой количества информации (на один символ), или плотностью энтропии Шеннона, поэтому более соответствует интуитивным представлениям о сложности. Стохастический ВР «несжимаем», поэтому имеет высокую алгоритмическую сложность, но система, которая его породила, может быть физически очень проста. С точки зрения физики важна сложность не самого ВР, а породившей его системы. Случаи, когда значения ВР полностью предсказуемы, соответствующие малой алгоритмической сложности, и когда они полностью непредсказуемы, соответствующие большой алгоритмической сложности, можно считать результатами измерений характеристик простых систем. Истинно сложные системы находятся между этими крайними случаями, что согласуется, например, с интуитивным представлением о сложности биологической системы и простотой идеального кристалла или стекла. В указанных крайних случаях энтропия либо постоянна, либо растет линейно с увеличением размера системы [19]. В обоих случаях поправки к асимптотическому поведению не увеличиваются при увеличении набора данных. Это дает возможность связать медленное приближение энтропии к асимптотическому пределу со сложностью системы [19, 20].

Из-за трудностей практического применения К-энтропии в случае ВР небольшой длины были предложены сначала приближенная энтропия [21], а затем выборочная энтропия [22], которые в пределе бесконечных ВР совпадают с К-энтропией, то есть тоже могут служить в качестве оценки сложности системы.

3. ПРИБЛИЖЕННАЯ ЭНТРОПИЯ И ВЫБОРОЧНАЯ ЭНТРОПИЯ

Приближенная (AppEn) и выборочная энтропии (SampEn) используются в качестве меры оценки случайности ВР в отсутствие сведений о системе, породившей этот ВР [23]. Этот подход основан на оценке неопределенности относительно возможных наборов данных наблюдений и связан, таким образом, с количеством информации, необходимой для полного их описания. Под неопределенностью понимают то, что возможно, но не известно. Все виды

энтропии были введены в качестве мер неопределенности относительно какого-либо события до того, как оно произошло.

AppEn вычисляется по следующему алгоритму [21, 23]. Для ряда $\{x_i\}_{i=1}^N$ строят m -мерные векторы в пространстве состояний (векторы вложения размерности m):

$$\mathbf{X}(i) = \{x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)\}, \\ 1 \leq i \leq N - m\tau.$$

Сходство векторов $\mathbf{X}(i)$ и $\mathbf{X}(j)$ определяют по расстоянию Чебышева:

$$d[\mathbf{X}(i), \mathbf{X}(j)] = d_{i,j} = \\ = \max_{0 \leq k \leq (m-1)} (|\mathbf{X}(i)_k - \mathbf{X}(j)_k|) = \\ = \max_{0 \leq k \leq (m-1)} (|x(i+k) - x(j+k)|).$$

Доля векторов $\mathbf{X}(j)$ на расстоянии не более r от $\mathbf{X}(i)$ составляет $C_i^{(m)}(r) = \frac{N_i^{(m)}(r)}{N - m\tau}$, где $N_i^{(m)}(r)$ — число индексов j , для которых $d_{i,j} \leq r$, $1 \leq j \leq N - m\tau$, затем вычисляется среднее значение логарифма этой доли векторов $\Phi^{(m)}(r) = \frac{1}{N - m\tau} \sum_{i=1}^{N - m\tau} \ln C_i^{(m)}(r)$. После этого размерность увеличивается на единицу и вычисляется $\Phi^{(m+1)}(r)$. Приближенная энтропия определяется следующим образом: $\text{AppEn}(m, r, \tau) = \Phi^{(m)}(r) - \Phi^{(m+1)}(r)$.

Выборочная энтропия вычисляется сходным образом, но исключаются сравнения векторов с самими собой [22, 23], из-за которого при вычислении AppEn происходит смещение в сторону большего сходства векторов. Доля векторов $\mathbf{X}_m(j)$ на расстоянии не более r от $\mathbf{X}_m(i)$ составляет $A_i^{(m)}(r) = \frac{N_i^{(m)}(r)}{N - m\tau - 1}$, $d_{i,j} \leq r$, $1 \leq j \leq N - m\tau$, $j \neq i$. Среднее значение по $1 \leq i \leq (N - m\tau)$ равно $\Psi^{(m)}(r) = \frac{1}{N - m\tau} \sum_{i=1}^{N - m\tau} A_i^{(m)}(r)$. Затем размерность увеличивается на единицу и вычисляется $\Psi^{(m+1)}(r)$: $m \rightarrow (m + 1) \rightarrow \Psi^{(m+1)}(r)$, после чего выборочная энтропия определяется по формуле: $\text{SampEn} = -\ln \frac{\Psi^{(m+1)}(r)}{\Psi^{(m)}(r)}$.

AppEn и SampEn были использованы для анализа электроэнцефалограмм (ЭЭГ) и магнитоэнцефалограмм при болезни Альцгеймера, SampEn применяли при анализе вариабельности частоты сердечных сокращений у крыс в разных ситуациях. Оба вида энтропии становятся неопределенными или приводят к неверным результатам при анализе коротких ВР, кроме того, критически зависят от выбора параметра r , поэтому неудобны в практических расчетах. Вычислительная сложность AppEn и SampEn квадратично растет с длиной ВР.

4. ПЕРЕСТАНОВОЧНАЯ ЭНТРОПИЯ И ЕЕ ВАРИАНТЫ

Перестановочная энтропия определяется как энтропия распределения упорядоченных отрезков ВР, которым ставится в соответствие определенный символ перестановки. Однако одно и то же распределение частот перестановок получается для разных абсолютных значений амплитуд ВР, поскольку в алгоритме учитываются только их относительные изменения. Это приводит к потере части исходной информации, содержащейся в ВР. Были предложены многочисленные модификации исходного алгоритма, имевшие целью учесть информацию, содержащуюся в изменениях амплитуд. Однако оценка сложности системы, характеризуемой данным ВР, не обязательно связана с абсолютными значениями амплитуд.

Перестановочная энтропия (РЕ) позволяет быстро количественно оценить неопределенность относительно временной структуры какого-либо ВР, полученного в результате измерений некоторой наблюдаемой величины, характеризующей изучаемую систему. Предложенный алгоритм не учитывает абсолютных значений амплитуд наблюдаемых величин, но только последовательность их относительных значений. При этом, конечно, теряется часть информации, содержащейся в ВР, но тем не менее РЕ во многих случаях позволяет определить, ведет себя система хаотически, стохастически или же закономерно. При увеличении длины ВР для любой хаотической системы РЕ стремится к энтропии Колмогорова–Синяя [24, 25].

Суть алгоритма состоит в следующем [26]. Скалярный ВР $\{x_i\}_{i=1}^N$ преобразуется в последовательность векторов (вложения размерности w):

$$\mathbf{X}_w(i) = (x_i, x_{i+\tau}, \dots, x_{i+(w-1)\tau}), \quad (1)$$

$$i = 1, 2, \dots, N - (w - 1)\tau.$$

Обычно используют $\tau = 1$ и $3 \leq w \leq 7$, при этом должно выполняться условие $N \gg w!$. Затем производится сортировка компонент каждого вектора: $x_{i+(j_1-1)\tau} \leq x_{i+(j_2-1)\tau} \leq \dots \leq x_{i+(j_w-1)\tau}$, где j_k ($k = 1, \dots, w$) — индекс компоненты вектора после сортировки (ранг). При равенстве компонент их порядок сохраняется: $x_{i+(j_{k1}-1)\tau} < x_{i+(j_{k2}-1)\tau}$, если $j_{k1} < j_{k2}$. Каждому вектору сопоставляется «символ» или «слово»: $X(i) \rightarrow A(i) = (j_1, \dots, j_{k1}, j_{k2}, \dots, j_w)$, при этом возможны $w!$ таких слов из w букв. Далее оценивают распределение вероятностей символов: $\{p_1, p_2, \dots, p_k\}$, $k \leq w!$ и рассчитывают перестановочную энтропию по Шеннону (например, в битах): $H_{PE}(w) = -\sum_{i=1}^k p_i \log_2 p_i$. Максимальное значение энтропии $H_{PE}(w) = \log_2(w!)$ соответствует равномерному распределению: $p_i = 1/w!$ и $k = w!$. Часто используют нормировку: $h_{PE} = H_{PE}(w)/\log_2(w!)$, при этом $0 \leq h_{PE} \leq 1$. Для регулярных и хаотических рядов $\lim_{w \rightarrow \infty} h_{PE} = 0$. Вычислительная сложность РЕ

растет линейно с увеличением числа членов ВР.

В исходном алгоритме распределение случайной величины предполагалось непрерывным, поэтому одинаковые значения компонент векторов должны встречаться редко и практически не должны влиять на результат. Авторы предложили для устранения проблемы равных значений добавлять малое случайное возмущение, но на практике это используется редко, потому что пригодно для непрерывных распределений или измерений с высоким разрешением. Было показано, что в случаях дискретных распределений, низкого разрешения или в исследованиях физиологических ВР — электрокардиограмм (ЭКГ), ЭЭГ и т.п. — исходный алгоритм может приводить к корреляциям, приводящим к ложным выводам [27]. С другой стороны, даже для ВР, полученного от полностью предсказуемой динамической системы, РЕ не обращается в нуль [28]. В связи с этим появились различные модификации и улучшения исходного алгоритма.

4.1. Модифицированная перестановочная энтропия (mPE)

В работе [29] предложено равным компонентам векторов присваивать один и тот же ранг. Так, например, двум векторам $\mathbf{X}_5(i) = (0.2, 0.5, 0.1, 0.2, 0.7)$, $\mathbf{X}_5(j) = (0.2, 0.5, 0.1, 0.24, 0.7)$ в алгоритме РЕ соответствовал бы один и тот же символ $A(i) = A(j) = (3, 1, 4, 2, 5)$. Модификация алгоритма: если $x_{i+(j_{k1}-1)\tau} = x_{i+(j_{k2}-1)\tau}$, а $j_{k1} < j_{k2}$, то обоим значения приписывается j_{k1} : $A'(i) = (j_1, \dots, j_{k1}, j_{k1}, \dots, j_w)$, то есть получим символы: $A'(i) = (3, 1, 1, 2, 5)$, $A'(j) = (3, 1, 4, 2, 5)$. С помощью частотной оценки определяется распределение вероятностей символов: $\{p'_1, p'_2, \dots, p'_k\}$, $k \leq k_w(w)$, где $k_w(w)$ — верхняя граница значений k . Для ее вычисления авторы использовали рекурсивный метод. Полученные значения для некоторых часто используемых значений w приведены в таблице. Значение mPE рассчитывается так же, как и РЕ: $H'_{PE}(w) = -\sum_{i=1}^k p'_i \ln p'_i$; нормировка: $h'_P = H'_{PE}(w)/\ln k_w(w)$, $0 \leq h'_{PE} \leq 1$.

Авторы [29] проанализировали ряды RR-интервалов в ЭКГ здоровых молодых, здоровых пожилых людей и пациентов с сердечной недостаточностью из открытых баз данных (MIT-BIH Fantasia database; BIDMC congestive heart failure database — входят в PhysioNet database, <https://physionet.org>) и показали, что РЕ не позволяет различать ЭКГ в этих группах, в то время как с помощью mPE группы различаются вполне надежно.

К недостаткам mPE относится очень большое значение $k_w(w)$ для $w \geq 4$, что накладывает ограничение на минимальную длину записи (таблица). Кроме того, mPE не имеет максимального значения для гауссовского белого шума, как это должно было бы быть. Так же, как РЕ, mPE сильно

зависит от уровня шума в анализируемом сигнале. В связи с этим была предложена «улучшенная» РЕ (IPE) [30].

Таблица. Верхние границы значений k для расчетов mPE

w	3	4	5	6	7
k_w	13	73	501	4051	37633

4.2. Улучшенная РЕ

Алгоритм расчета IPE позволяет до некоторой степени компенсировать указанные выше недостатки и, кроме того, учесть амплитуды сигналов. После построения векторов (1) производят однородную дискретизацию (UQ) первого столбца $\mathbf{X}(:, 1)$ матрицы, образованной векторами-строками:

$$UQ(\mu) = \begin{cases} 0, & x_{min} \leq x < x_{min} + \Delta, \\ 1, & x_{min} + \Delta \leq x < x_{min} + 2\Delta, \\ \vdots & \\ L-1, & x_{min} + (L-1)\Delta \leq x \leq x_{max}, \end{cases}$$

где x_{min} и x_{max} — минимальное и максимальное значения исследуемого ВР, μ — входные данные, $\Delta = (x_{max} - x_{min})/L$, L — уровень дискретизации. Величина UQ — целое число в диапазоне от 0 до $L-1$ — «символ» $\mathbf{X}(:, 1)$. Обозначим его $S(:, 1)$. Для k -го столбца $\mathbf{X}(:, k)$, $2 \leq k \leq w$ значение $S(:, k)$ рассчитывается по формуле:

$$S(j, k) = S(j, 1) + \left[\frac{X(j, k) - X(j, 1)}{\Delta} \right], \\ 1 \leq j \leq N - (w-1)\tau, \quad 2 \leq k \leq w.$$

Строки полученной матрицы рассматриваются как «слова» $A(l)$, $1 \leq l \leq L^w$. После этого рассчитывают оценку распределения вероятностей этих слов $\{p_l\}$ и нормированное значение IPE:

$$H_{IPE}(w, \tau, L) = \frac{-\sum_{l=1}^h p_l \ln p_l}{\ln L^w},$$

где $h \leq L^w$, а $\ln L^w$ — максимальное значение H_{IPE} , которое достигается только для равномерного распределения.

Основные отличия IPE от РЕ в следующем. Во-первых, учитывается информация об амплитудах и их флуктуациях, во-вторых, равным значениям приписываются одинаковые символы, в-третьих, IPE более устойчива по отношению к шуму и последнее — количество возможных «слов» в IPE составляет L^w , а не $w!$.

Авторы проанализировали влияние параметров на оценки IPE и рекомендуют для большинства практических ситуаций использовать $w = 4$, $\tau = 1$, $L = 4$. Анализ синтезированных и естественных ВР показал, что IPE существенно превосходит по чувствительности РЕ и mPE.

4.3. Усредненная РЕ (Ensemble PE, EPE)

Важным параметром при вычислении РЕ является размерность w_j векторов $\mathbf{X}_{w_j}(i)$. Обычно требуется, чтобы выполнялось условие $(w_{max} + 1)! \leq N$, то есть $2 \leq w_j \leq w_{max}$. В работе [31] для уменьшения зависимости РЕ от выбора w_j предложено использовать значение РЕ, усредненное по всем допустимым значениям w_j :

$$h_{PE}(\mathbf{X}_{w_j}, w_j) = H_{PE}(\mathbf{X}_{w_j}, w_j) / \log_2(w_j!); \\ h_{EPE} = \frac{\sum_{j=2}^{w_{max}} h_{PE}(\mathbf{X}_{w_j}, w_j)}{w_{max} - 1}.$$

Авторы на разных примерах показывают, что применение такого подхода уменьшают зависимость результатов расчетов от шума и обычно лучше выявляют различия между ВР.

4.4. Усредненная IPE (EIPE)

В работе [32] в отличие от исходного алгоритма расчета IPE авторы предлагают сначала провести нормировку исходных значений ВР $\{x_i\}_{i=1}^N$ с помощью нормальной функции распределения. При этом новые члены ВР будут иметь следующие значения:

$$y_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_i} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt, \quad (2)$$

где μ и σ^2 — среднее значение и дисперсия исходного ВР. Затем выбираются ширина окна w и шаг по времени τ и строится «фазовое пространство»: $\Upsilon(j, :) = [y_j, y_{j+\tau}, \dots, y_{j+(w-1)\tau}]$, где $\Upsilon(j, :)$ — j -я строка матрицы и $j = 1, 2, \dots, N - (w-1)\tau$.

Обозначим y_{min} и y_{max} минимальное и максимальное значения членов ВР. Затем надо выбрать число уровней дискретизации L , определить шаг дискретизации $\Delta = (y_{max} - y_{min})/L$ и построить равномерную функцию распределения $UPF(u)$, где $u \in (y_{min}, y_{max})$:

$$UPF(u) = \begin{cases} 0 & \text{при } y_{min} \leq u < y_{min} + \Delta, \\ 1 & \text{при } y_{min} + \Delta \leq u < y_{min} + 2\Delta, \\ \vdots & \\ L-1 & \text{при } y_{max} - \Delta < u \leq y_{max}. \end{cases}$$

Таким образом, UPF преобразует u в целочисленный символ из диапазона от 0 до $L-1$. Применив это преобразование к элементам первого столбца $\Upsilon(:, 1)$, получим столбец символов $S(:, 1)$. Для k -го столбца ($2 \leq k \leq w$) используется преобразование:

$$S(j, k) = S(j, 1) + \lfloor [\Upsilon(j, k) - \Upsilon(j, 1)] / \Delta \rfloor,$$

где $1 \leq j \leq \lfloor [N - (w-1)\tau] \rfloor$, а $\lfloor \cdot \rfloor$ обозначает функцию округления вниз к ближайшему целому числу. В результате получается матрица S , каждая строка которой считается символом (или словом). Для

вычисления IPE надо определить вероятности этих символов p_i . Так как каждый символ содержит w элементов, которые могут принимать L различных значений, то полное число символов равно L^w . IPE вычисляется по формуле Шеннона и нормируется на $\ln L^w$:

$$H_{IPE}(w, L, \tau) = - \sum_{i=1}^{L^w} p_i \ln p_i / \ln L^w.$$

Многие свойства IPE сильно зависят от величины L . Чтобы до некоторой степени избавиться от этой зависимости, авторы предлагают провести усреднение по нескольким значениям L :

$$H_{EIPE}(w, \tau) = \frac{1}{L_{max} - L_{min}} \sum_{L=L_{min}}^{L_{max}} H_{IPE}(w, L, \tau),$$

где L_{min} и L_{max} — выбранные минимальное и максимальное значения L . При малых значениях L увеличивается устойчивость по отношению к шуму, но теряется часть информации, содержащейся в ВР.

Сравнение результатов применения разных видов энтропии (PE, взвешенная PE, дисперсионная энтропия, которая будет рассмотрена в разд. 4.3, и EIPE) к анализу синтезированных шумовых рядов (белый, розовый и коричневый) при $20 \leq N \leq 700$, $w = 4$, $\tau = 1$ и $2 \leq L \leq 8$ показало, что во всех случаях белый шум имеет максимальную энтропию, а коричневый — минимальную, но EIPE позволяет надежно различать вид шума даже при минимальных длинах ВР, когда другие виды энтропии этого делать не могут.

Анализ логистического отображения $x_{n+1} = r x_n (1 - x_n)$ при изменении r от 3.5 до 3.99 с шагом 0.001 и при $N = 10000$ показал,

$$p(u_{v_0 v_1 \dots v_{w-1}}) = \frac{\text{Number} \{t | t \leq N - (w-1)\tau, \text{ когда } \mathbf{Z}_{w,c}(i) \text{ имеет паттерн } u_{v_0 v_1 \dots v_{w-1}}\}}{N - (w-1)\tau},$$

где $\text{Number} \{ \cdot \}$ — функция подсчета числа событий.

Для пояснения последовательности шагов вычислений рассмотрим конкретный пример. Пусть $\{x_i\}_{i=1}^N = (9, 8, 1, 12, 5, -3, 1.5, 8.01, 2.99)$. После отображения на интервал $[0, 1]$ получим $\{y_i\}_{i=1}^N = (0, 82, 0.75, 0.21, 0.94, 0.52, 0.05, 0.24, 0.75, 0.35)$. Округление при $c = 3$ дает $\{z_i\}_{i=1}^N = (3, 3, 1, 3, 2, 1, 1, 3, 2)$. Векторы вложения при $\tau = 1$ и $w = 2$ имеют вид: $\mathbf{Z}_{2,3}(1) = (3, 3)$, $\mathbf{Z}_{2,3}(2) = (3, 1)$, $\mathbf{Z}_{2,3}(3) = (1, 3), \dots, \mathbf{Z}_{2,3}(8) = (3, 2)$ и им соответствуют паттерны дисперсии $u_{33}, u_{31}, u_{13}, \dots, u_{32}$, а число возможных паттернов (индексов) равно $3^2 = 9$. Поэтому вероятности паттернов будут следующими: $p(u_{11}) = \frac{1}{8}$, $p(u_{12}) = 0$, $p(u_{13}) = \frac{1}{4}$, $p(u_{21}) = \frac{1}{8}, \dots, p(u_{33}) = \frac{1}{8}$.

Дисперсионная энтропия вычисляется по формуле Шеннона:

$$DE(\mathbf{X}, w, c, \tau) = - \sum p(u_{v_0 v_1 \dots v_{w-1}}) \ln p(u_{v_0 v_1 \dots v_{w-1}}),$$

что, начиная со значения $r = 3.57$, наблюдается положительная корреляция EIPE с увеличением хаотичности ряда, а при значениях, соответствующих периодическим режимам (в частности, вблизи 3.84), наблюдается резкое уменьшение EIPE. Этот алгоритм также показал несколько лучшие результаты при анализе естественных сигналов, таких как ЭКГ, ЭЭГ, вибрации исправных и неисправных подшипников, а также морских судов.

4.5. Дисперсионная энтропия (DE)

Среди других видов энтропии хорошую способность к различению типов ВР, устойчивость к шумам, способность к учету амплитуд имеет дисперсионная энтропия [33]. Ее вычисление состоит из следующих этапов.

Значения элементов ряда $\{x_i\}_{i=1}^N$ как-либо отображаются на интервал $[0, 1]$, обычно с помощью нормальной кумулятивной функции распределения (2), в результате получается последовательности $\{y_i\}_{i=1}^N$. Члены y_i нового ряда распределяются по c классам — заменяются целыми числами из интервала $[1, c]$, при этом авторы рекомендуют использовать функцию округления: $r_c(j) = c \cdot y_j + 0.5$, $z_c(j) = \text{round}(r_c(j))$, $j = 1, 2, \dots, N$, а число классов выбирать из интервала [3, 9] так, чтобы $c^w < N$. Для размерности вложения w и временной задержки τ строят векторы вложения $\mathbf{Z}_{w,c}(i) = \{z_c(i), z_c(i + \tau), \dots, z_c(i + (w-1)\tau)\}$, $i = 1, 2, \dots, N - (w-1)\tau$. Каждому вектору $\mathbf{Z}_{w,c}(i)$ сопоставляется «паттерн дисперсии» $u_{v_0 v_1 \dots v_{w-1}}$, при этом $v_0 = z_c(i)$, $v_1 = z_c(i + \tau), \dots, v_{w-1} = z_c(i + (w-1)\tau)$. Число таких паттернов равно c^w . Затем оценивается вероятность появления каждого паттерна:

и обычно нормируется на максимальное возможное значение DE (когда все паттерны равновероятны):

$$\text{NDE}(\mathbf{X}, w, c, \tau) = \frac{\text{DispEn}(\mathbf{X}, w, c, \tau)}{\ln c^w}.$$

В приведенном примере $DE \approx 1.844$.

Для реальных ВР отнесение членов ряда к классам с помощью функции округления не всегда однозначно, поэтому в работе [34] предложено заменить функцию округления на некую «нечеткую» (fuzzy) функцию, в результате чего член ВР может быть отнесен к двум классам одновременно и каждому вложению может соответствовать не более 2^w паттернов. Это не приводит к увеличению вычислительной сложности алгоритма (как и для DE, она растет линейно с длиной ВР), но позволяет использовать более короткие ВР.

Еще одна модификация DE — кодированная DE

(CDE) — предложена в работе [35]. Операция кодирования была использована ранее для улучшения свойств PE [36], но другие модификации PE давали лучшие результаты. В применении к DE она сводится к дополнительному квадратичному разбиению вложений $\mathbf{Z}_{w,c}(i)$ и также позволяет устранить недостатки использования функции округления. Это осуществляется следующим образом.

Каждому вложению $\mathbf{Z}_{w,c}(i)$, имеющему один и тот же паттерн дисперсии, сопоставляется вложение $\mathbf{X}_w(i)$ исходного ряда, и элементы вложения $\mathbf{Z}_{w,c}(i)$ заменяются на средние значения соответствующих элементов $\mathbf{X}_w(i)$. Затем формируется средний паттерн:

$$\langle u_{v_1 \dots v_w} \rangle = \{ \langle z_{v_1 \dots v_w}^1 \rangle, \langle z_{v_1 \dots v_w}^2 \rangle, \dots, \langle z_{v_1 \dots v_w}^w \rangle \},$$

где $\langle z_{v_1 \dots v_w}^i \rangle$ — среднее значение элемента вложения, а $\langle u_{v_1 \dots v_w} \rangle$ — общий паттерн. Так, например, для рассмотренного примера исходные вложения имеют вид $\mathbf{X}_2(1) = (9, 8)$, $\mathbf{X}_2(2) = (8, 1)$, \dots , $\mathbf{X}_2(8) = (8.01, 2.99)$. К паттерну u_{13} относятся вложения $\mathbf{Z}_{2,3}(2)$ и $\mathbf{Z}_{2,3}(7)$, которым соответствуют вложения $\mathbf{X}_2(2) = (8, 1)$ и $\mathbf{X}_2(7) = (1.5, 8.01)$. Тогда средние значения будут равны $\langle z_{12}^1 \rangle = 4.75$, $\langle z_{12}^2 \rangle = 4.005$, то есть средний паттерн $\langle u_{13} \rangle = \{4.75, 4.005\}$. В работе [35] предложено сравнивать «неокругленные» значения $r_c(i)$ со средним значением и производить разбиение по следующему критерию:

$$v(i) = \begin{cases} 2, & r_c(i) > \langle z_{v_1 \dots v_w}^i \rangle, \\ 1, & r_c(i) = \langle z_{v_1 \dots v_w}^i \rangle, \\ 0, & r_c(i) < \langle z_{v_1 \dots v_w}^i \rangle. \end{cases}$$

Но на самом деле сравнение по этому критерию проводилось со значениями вложений исходного ряда, то есть $\langle u_{13} \rangle$ поэлементно сравнивали с $\mathbf{X}_2(2) = (8, 1)$ и $\mathbf{X}_2(7) = (1.5, 8.01)$ и применяли указанный критерий. В нашем случае получим $v(2) = (2, 0)$, $v(7) = (0, 2)$.

В результате получается вектор $\mathbf{v}_{w,c}(i) = \{v(i), v(i + \tau), \dots, v(i + (w - 1)\tau)\}$. Объединив оба разбиения (на классы и последнее), получим «объединенный» вектор $\{\mathbf{Z}_{w,c}(i), \mathbf{v}_{w,c}(i)\}$, которому соответствует паттерн $u_{v_1 v_2, \dots, v_{2w}}$, при этом число возможных паттернов возрастает до $c^w \cdot 3^w$, поскольку в критерии имеются три возможности.

На практике значения $v(i) = 1$ встречаются очень редко, поэтому без ущерба для точности такие случаи можно не учитывать, что приводит к упрощенному критерию второго разбиения:

$$v(i) = \begin{cases} 1, & r_c(i) > \langle z_{v_1 \dots v_w}^i \rangle, \\ 0, & r_c(i) \leq \langle z_{v_1 \dots v_w}^i \rangle, \end{cases}$$

при этом число возможных паттернов уменьшается до $c^w \cdot 2^w$. Расчет упрощенной CDE (SCDE) проводится по обычной схеме: оцениваются вероятности паттернов:

$$p(u_{v_1 v_2 \dots v_{2w}}) = \frac{\text{Number}\{u_{v_1 v_2 \dots v_{2w}}\}}{N - (w - 1)\tau},$$

затем используется формула Шеннона:

$$SCDE = - \sum_{u=1}^{c^w \cdot 2^w} p(u_{v_1 v_2 \dots v_{2w}}) \ln p(u_{v_1 v_2 \dots v_{2w}})$$

и при необходимости производится нормировка:

$$NSCDE = \frac{SCDE}{\ln(c^w \cdot 2^w)}.$$

Результаты расчетов CDE и SCDE для синтезированных и естественных сигналов (ЭЭГ, вибрации подшипников, шумы морских судов) не менее, чем на 10% превосходят другие виды энтропии (PE, CPE, DE) в точности распознавания, но для расчета SCDE требуется примерно в 3 раза меньше времени, чем для расчета CDE.

ЗАКЛЮЧЕНИЕ

В настоящем кратком обзоре рассмотрены только базовые сведения о возможностях применения энтропийных методов в анализе ВР, которые могут быть интересны экспериментаторам и аналитикам big-data. Основное внимание уделено перестановочной энтропии и ее модификациям, поскольку эти виды энтропии широко используются в различных приложениях. В обзор не вошли многие другие методы, разработанные на основе рассмотренных подходов (пузырьковая, нечеткая, многовариантная, многомасштабная и другие виды энтропии), а также методы с использованием, например, неэкстенсивной энтропии Тсаллиса, хотя они, возможно, могут в некоторых случаях быть более эффективными (см., например, [37–53]).

В целом перестановочная энтропия часто дает хорошие результаты при классификации ВР, особенно достаточно длинных. В случае очень коротких ВР или при необходимости обнаружить кратковременное воздействие каких-либо факторов на систему лучше использовать другие виды энтропии, рассмотренные выше. Обычно для сокращения времени расчетов используют малые значения параметров w и L , но иногда имеет смысл провести дополнительный анализ зависимости результатов расчетов от этих параметров.

Программное обеспечение для всех описанных методов имеется в открытом доступе в Интернете (см., например, [54]).

- [1] Генералов Е.А., Левашова Н.Т., Сидорова А.Э. и др. // Биофизика. **62**, № 5. 660 (2017). (*Generalov E.A., Levashova N.T., Sidorova A.E. et al. // Biophysics. 62, N 5, 717 (2017)*).
- [2] Gaidin S., Maiorov S., Laryushkin D. et al. // *J. Neurochem. 164, № 5. 583 (2023)*.
- [3] Кульченко Н.Г., Яценко Е.В. // Экспериментальная и клиническая урология. **11**, № 3. 158 (2019).
- [4] Генералов Е.А., Симоненко Е.Ю., Кульченко Н.Г., Яковенко Л.В. // Биомедицинская химия. **68**, № 6. 403 (2022).
- [5] Generalov E., Clarke T., Iddamalgoda L. et al. *Systems Biology in Biomarker Development in Cancer Signaling Therapy*. Elsevier book. Companion and Complementary Diagnostics: From Biomarker Discovery to Clinical Implementation, 2019.
- [6] Kritskaya K.A., Fedotova E.I., Berezhnov A.I. // *Biomedicines. 12, N 2, 282 (2024)*.
- [7] Whitney H. // *Ann. Math. 37*, 645 (1936).
- [8] Packard N.H., Crutchfield J.P., Farmer J.D., Shaw R.S. // *Phys. Rev. Lett. 45, N 9, 712 (1980)*.
- [9] Roux J.-C., Simoyi R.H., Swinney H.L. // *Physica D. 8*, 257 (1983).
- [10] Shannon C.E. // *Bell Syst. Tech. J. 27, N 3. 379 (1948)*.
- [11] Shannon C.E. // [10.1002/j.1538-7305.1948.tb00917.x](https://doi.org/10.1002/j.1538-7305.1948.tb00917.x) *Bell Syst. Tech. J. 27, N 4. 623 (1948)*.
- [12] Чернавский Д.С., Хазин М.Л. // Краткие сообщения по физике ФИАН. № 10. 39 (2001).
- [13] Jaynes E.T. // *Phys. Rev. 106, N 4, 620 (1957)*.
- [14] Jaynes E.T. // *Phys. Rev. 108, N 2, 171 (1957)*.
- [15] Хинчин А.Я. // УМН. **8**, N 3, 3 (1953).
- [16] Хинчин А.Я. // УМН. **11**, N 1, 17 (1956).
- [17] Bialek W.S. *Biophysics: searching for principles*. Princeton, UK: Princeton University Press, 2012; ISBN 978-0-691-13891-6.
- [18] Grassberger P., Procaccia I. // *Phys. Rev. A. 28, N 4, 2591 (1983)*.
- [19] Bialek W., Nemenman I., Tishby N. // *Neural Comput. 13*, 2409 (2001).
- [20] Grassberger P. // *Int. J. Theor. Phys. 25*, 907 (1986). [10.1007/BF00668821](https://doi.org/10.1007/BF00668821)
- [21] Pincus S.M. // *Proc. Natl. Acad. Sci. 88, N 6, 2297 (1991)*.
- [22] Richman J.S., Moorman J.R. // *Am. J. Physiol. Heart. Circ. Physiol. 278*, N 6, 2039(2000).
- [23] Delgado-Bonal A., Marshak A. // *Entropy. 21*, 541 (2019).
- [24] Rosso O.A., Larrondo H.A., Martin M.T. et al. // *Phys. Rev. Lett. 99*, 154102 (2007).
- [25] Keller K., Mangold T., Stolz I., Werner J. // *Entropy 19*. 134 (2017).
- [26] Bandt C, Pompe B. // *Phys Rev Lett. 88*, 174102 (2002).
- [27] Zunino L., Olivares F., Scholkmann F., Rosso O.A. // *Phys. Lett. A. 381*, N 22, 1883 (2017).
- [28] Fouda J.S.A.E., Koepf W. // *Commun. Nonlinear Sci. Numer. Simulat. 27*, 216 (2015).
- [29] Bian C., Qin C., Ma Q.D.Y., Shen Q. // *Phys. Rev. E. 85*, 021906 (2012). [10.1103/PhysRevE.85.021906](https://doi.org/10.1103/PhysRevE.85.021906)
- [30] Chen Z., Li Y., Liang H., Yu J. // *Complexity. Article ID 1403829 (2019)*.
- [31] Azami H., Sanei S., Rajji T.K. // *Knowledge-Based Systems. 256*, 109876 (2022). [10.1016/j.knosys.2022.109876](https://doi.org/10.1016/j.knosys.2022.109876)
- [32] Chen Z., Ma X., Fu J., Li Y. // *Entropy. 25*, 1175 (2023).
- [33] Rostaghi M., Azami H. // *IEEE Signal Processing Lett. 23*, N 5, 610 (2016). [10.1109/LSP.2016.2542881](https://doi.org/10.1109/LSP.2016.2542881)
- [34] Rostaghi M., Khatibi M.M., Ashory M.R., Azami H. // *IEEE Transactions on Fuzzy Syst. 30*, N 9, 3785 (2022).
- [35] Li Y., Geng B., Tang B. // *Nonlinear Dyn. 111*, 9327 (2023).
- [36] Kang H., Zhang X., Zhang G. // *Entropy. 22*, N 2. 187 (2020).
- [37] Chen W., Wang Zh., Xie H., Yu W. // *IEEE Transact. Neural Syst. Rehabil. Eng. 15*, N 2 (2007).
- [38] Staniek M., Lehnertz K. // *Phys. Rev. Lett. 100*, 158101 (2008).
- [39] Bose R., Chouhan S. // *Phys. Rev. E. 83*, 051918 (2011).
- [40] Hanel R., Thurner S., Gell-Mann M. // *PNAS. 108*, N 16, 6390 (2011).
- [41] Lobier M., Siebenhuehner F., Palva S., Palva J.M. // *NeuroImage. 85*, 853 (2014). [10.1016/j.neuroimage.2013.08.056](https://doi.org/10.1016/j.neuroimage.2013.08.056)
- [42] Li P., Li K., Liu C. et al. // *IEEE Transact. Biomed. Eng. 63*, N 11 (2016).
- [43] Tempesta P. // *Proc. R. Soc. A. 472*, 20160143 (2016).
- [44] Hsu C.F., Wei S.-Y., Huang H.-P., Hsu L. et al. // *Entropy. 19*, 550 (2017).
- [45] Manis G., Aktaruzzaman M., Sassi R. // *IEEE Transact. Biomed. Eng. 64*, N 11, 2711 (2017).
- [46] Li Y., Gao X., Wang L. // *Sensors. 19*, 5203 (2019).
- [47] Azami H., da Silva L.E.V., Omoto A.C.M., Humeau-Heurtier A. // *Signal Processing: Image Commun. 75*, 178 (2019).
- [48] Qin G., Shang P. // *Int. J. Bifurc. Chaos. 31*, N 9, 2150128 (2021).
- [49] Deka B., Deka D. // *Chaos, Solitons and Fractals 158*, 112101 (2022).
- [50] Jiang X., Yi Y., Wu J. // *Front. Phys. 11*, 1163767 (2023).
- [51] Ji G. // *Front. Phys. 11*, 1146493 (2023). [10.3389/fphy.2023.1146493](https://doi.org/10.3389/fphy.2023.1146493)
- [52] Tsallis C. // *J. Stat. Phys., 52*, 479 (1988).
- [53] Li C., Shang P. // *Physica A. 523*, 10 (2019). [10.1016/j.physa.2019.01.031](https://doi.org/10.1016/j.physa.2019.01.031)
- [54] Flood M.W., Grimm B. // *PLoS ONE. 16*, N 11, e0259448 (2021). www.EntropyHub.xyz.

Using Entropy in Time Series Analysis

A.M. Adelyanov^a, E.A. Generalov, Weng Chjeng, L.V. Yakovenko

¹Faculty of Physics, Lomonosov Moscow State University
Moscow 119991, Russia

E-mail: ^aleo.yakovenko@mail.ru

Results of experiments in the field of biophysics are often presented as time series obtained with low resolution and not always of great length. In particular, in studies of the effects of various physicochemical factors on bilayer lipid membranes, transmembrane ion currents and their fluctuations are usually measured. In this case, the mean values and variances of the currents may not differ significantly, making it difficult to determine the nature and degree of impact based on them. Therefore, the development of approaches to time series analysis has never ceased. Attempts to use the entropy of random variable distributions in such analysis have been made for a long time, but in practical work, these approaches have been difficult to implement, especially due to the requirements for the length of the series and the absence of noise. In recent decades, there have been significant changes in this area, and many new methods of time series analysis using various modifications of entropy have been proposed. In this regard, there is a need for a summary of methods based on entropy calculation, indicating their advantages and disadvantages. This is the goal of the proposed brief review of entropy-based methods for analyzing scalar time series, which can be useful in analyzing experimental data. The review considers only some of the basic approaches on which further algorithmic improvements are based. The concept of entropy sometimes causes difficulties for students, so the review can also be useful for educational purposes.

PACS: 02.50.-r; 02.50.Fz; 01.50.-i

Keywords: time series, entropy, complexity.

Received 13 March 2024.

English version: *Moscow University Physics Bulletin*. 2024. **79**, No. 4. Pp. 415–425.

Сведения об авторах

1. Адельянов Артем Маратович — ведущий программист; e-mail: tashxat@gmail.com.
2. Генералов Евгений Александрович — канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: generals1179@gmail.com.
3. Вэнь Чжэнь — студент; e-mail: ven.c19@physics.msu.ru.
4. Яковенко Леонид Владимирович — доктор физ.-мат. наук, профессор; e-mail: leo.yakovenko@mail.ru.