

МАТЕМАТИЧЕСКАЯ И ФИЗИЧЕСКАЯ ПРИРОДА ПРОПУСКНОЙ СПОСОБНОСТИ

Введение и постановка проблемы

В настоящее время самыми употребляемыми категориями в теории связи и ее приложениях стали определения фундаментальных пределов скорости и надежности передачи информации, пропускной способности каналов и величины отношения сигнал/шум как определяющего показателя прогнозируемого качества связи. Работы В. А. Котельникова [1] и К. Шеннона [2], изданные в 1946 – 1948 гг., считаются открытием фундаментальных законов сжатия, передачи и приема данных и знаменуют рождение теории передачи информации в ее современном понимании. Теория, основанная на глубоком пересечении с теорией вероятностей, статистикой, информатикой и другими областями знаний, явилась основой для развития коммуникаций, хранения и обработки данных и других информационных технологий. Эту теорию можно определить как науку, занимающуюся исследованием и оптимизацией алгоритмов кодирования и декодирования информации в целях создания экономных и надежных способов ее передачи по каналам связи и хранения в запоминающих устройствах. Она возникла из потребностей радиосвязи, радиолокации, телефонии, телевидения, вычислительной техники и составляет теоретический фундамент для построения систем связи. В центре внимания этой теории проблема оптимального (с точки зрения скорости, надежности и эффективности) использования имеющихся технических средств для передачи, преобразования, распределения и хранения информации. В настоящее время по степени глубины и объему выполненных исследований теория информации сравнялась со многими разделами математической физики.

Несомненно, основной категорией современной теории информации является понятие пропускной способности зашумленного канала, которая была определена Шенноном [2, 6]. В его трактовке пропускная способность – это граница скорости передачи информации, которую нельзя превзойти при сколь угодно высокой надежности передачи *ни при каких способах кодирования и декодирования*, но к которой, в то же время, можно сколь угодно точно приблизиться за счет их соответствующего выбора. Пропускная способность канала была выражена в статистических терминах через введенную математическую характеристику совместного распределения вероятностей двух случайных величин, названную количеством информации. Она равна максимуму количества информации в сигнале на выходе канала относительно сигнала на его входе, где максимум берется по всем возможным распределениям вероятностей сигнала на входе. Количество информации, в свою очередь, выражается через другую величину, издавна использовавшуюся в термодинамике, – энтропию и представляет собой разность между энтропией сигнала на выходе канала и его условной энтропией, если известен сигнал на входе. Методологическая роль пропускной способности в теории информации чрезвычайно высока, поскольку она является не только основой для сформулированной Шенноном теоремы кодирования, но и инструментом для доказательства большинства других фундаментальных теорем и существующих пределов.

Несмотря на неоспоримые достижения в теории информации, она, в последнее время все чаще, подвергается критике. Основанием для этого является, в определенной мере, отсутствие практического конструктивизма в различных формулировках теорем, но и все более явно проявляющиеся проблемы развития. Видимый технологический прогресс всех видов сервисов связи практически не способен скрыть факт отсутствия существенного прироста показателей удельной эффективности телекоммуникационного оборудования. Протоколы

канального и физического уровня систем передачи информации (СПИ) являются, в значительной мере, затратными. Помехоустойчивые коды, история теоретических и экспериментальных исследований которых уже насчитывает более 70 лет, в режиме коррекции ошибок почти не используются на практике. Причиной тому является не только вычислительная сложность построения и декодирования громоздких конструкций в скоростных каналах, но и неприемлемость довольно существенной остаточной величины вероятности декодирования (исправления) с ошибкой при передаче данных и текстов компьютерных программ. Без большого преувеличения можно сказать, что удельная эффективность телекоммуникаций сохранилась на уровне 20-х годов прошлого столетия. Развитие техники и технологий связи носит сугубо экстенсивный характер. Приращение производительности достигается практически исключительно за счет совершенствования технологической базы реализации приемопередающих устройств, а также расширения полосы частот и мощности передатчиков (что, собственно, и предопределяет математическое определение пропускной способности). Это имеет определенные негативные последствия в моральном, материальном и экологическом аспектах. Все более существенной становится проблема электромагнитной совместимости. Перегруженные традиционные диапазоны радиочастот и малая полоса пропускания металлических линий связи вынудили перейти на использование оптического диапазона, возможности которого также не безграничны, но потребовали изменения (удорожания и снижения надежности) дополнительных устройств согласования сигналов. Мобильные технологии, начиная от 2G, LTE и до перспективной 5G, не претерпевают кардинальных изменений, за исключением экстенсивных доработок. Геостационарная орбита спутников связи приближается к предельному насыщению космическими объектами. Скорость возрастания потребностей в услугах связи превышает скорость приращения производительности СПИ. Все это, несомненно, является свидетельством явных проблем теории и практики построения систем передачи информации.

Целью данной работы является выявление трех основных причин – устаревших положений, заложенных в «фундамент» теории информации и являющихся первопричиной ее тупиковой эволюции, в частности:

- доказательство очевидных методологических неточностей в существующих определениях пропускной способности непрерывных каналов СПИ;
- обоснование некорректности обобщающей канонизации пропускной способности как предела достижимых скоростей для любых моделей непрерывных каналов;
- преувеличение роли принятия решений на основе правила максимального правдоподобия как наилучшего способа оценки состояния выхода зашумленного канала при малых отношениях сигнал/шум.

Несомненно, рассматриваемые вопросы могут выглядеть в определенной мере дискуссионными, особенно в отрыве от контекста полученных новых научных результатов, не являющихся предметом данной работы. Основное содержание статьи следует рассматривать как комплекс мотиваций, побуждающих к поиску принципиально новых решений в математической теории связи, соответствующей истинному физическому содержанию процесса передачи информации.

1. Дифференциальная энтропия непрерывных распределений и аналитическое определение пропускной способности гауссова канала

В начале рассмотрения предмета работы придется обратиться к некоторым общеизвестным фактам. Первое определение пропускной способности дискретного двоичного канала без памяти с симметричным графом переходов, задаваемым диагональной вероятностью ошибки p_0 дано в [2] и использует статистическую меру неопределенности дискретного выбора, названную энтропией:

$$C = V \cdot \max_{P(X)} \{H(X) - H(X|Y)\}, \quad (1)$$

где X, Y – сообщения на входе и выходе зашумленного канала; $P(X) = \{p(0), p(1)\}$ – распределение вероятностей появления символов двоичного алфавита; V – количество двоичных символов, передаваемых по каналу за 1 секунду;

$$H(X) = -\{p(0)\log[p(0)] + p(1)\log[p(1)]\} \quad (2)$$

– энтропия (неопределенность) источника двоичных сообщений (при измерении информации в битах – основание логарифма равно двум);

$$H(X|Y) = -\{p_0 \log[p_0] + (1-p_0)\log[1-p_0]\} \quad (3)$$

– ненадежность канала – энтропия (неопределенность) помехи. При известном качестве канала, задаваемом параметром p_0 , максимум (1) достигается при равновероятных символах источника $p(0) = p(1) = 1/2$ и составляет величину:

$$C = V[1 - H(X|Y)]. \quad (4)$$

Определение $I_b = C/V$ часто используется для вычисления среднего количества информации, содержащейся в одном двоичном символе на выходе зашумленного двоичного канала, в частности при оценке показателей удельной эффективности СПИ [3].

Формула (4) получила обобщение для случая недвоичного канала без памяти (см., например, [4, 10]). К настоящему времени, кроме указанных случаев, для моделей дискретных каналов известны аналитические определения для пропускной способности канала со стиранием и некоторые «экзотические» примеры с несимметричными графами переходов, рассмотренные самим К. Шенноном в оригинальной работе [2].

Сама по себе модель любого дискретного канала является своего рода надстройкой над моделью непрерывного (по времени и уровню) канала. Формулы (1) – (4) объективно понятны, имеют прозрачный физический и математический смысл и не будут обсуждаться в дальнейшем. Их рассмотрение в данной работе необходимо для отслеживания преемственности методологического подхода, использованного Шенноном при аналитическом выводе формулы пропускной способности непрерывного канала. В существующей парадигме класс непрерывных каналов, для которых определена пропускная способность, сужен до «гауссовых» [2, 5, 7, 8]. Ниже будет показано, что это неверно.

Для непрерывного источника, сообщения которого выбираются из бесконечного множества, Шеннон, следуя логике (1) – (4), вводит понятие энтропии непрерывного распределения (которую часто называют дифференциальной энтропией):

$$H(X) = - \int_{-\infty}^{\infty} f(x) \log[f(x)] dx, \quad (5)$$

где $f(x)$ – функция плотности распределения вероятностей (ПРВ) непрерывной случайной величины x . Соответственно, совместная и условная энтропия двух статистически связанных случайных аргументов, определяющих вход и выход непрерывного канала, задаются выражениями:

$$H(X, Y) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log[f(x, y)] dx dy; \quad (6)$$

$$H(Y|X) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log \left[\frac{f(x, y)}{f(y)} \right] dx dy. \quad (7)$$

К основным свойствам энтропии непрерывного случая (5) относятся следующие:

1) при заданном ограничении на среднюю мощность σ^2 непрерывного, центрированного относительно нуля процесса, энтропия (5) максимальна в случае, если этот процесс гауссов, т.е.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (8)$$

при этом

$$\max_{f(x)} (H(x)) = \log \sqrt{2\pi e \sigma^2}; \quad (9)$$

2) в отличие от дискретного случая, в определениях Шеннона (5) – (7), (9) [см. 2] измерение дифференциальной энтропии производится относительно заданной системы координат, т.е. не является абсолютным. Это означает, что в случае, когда аргумент логарифма после вычисления интегралов меньше единицы, дифференциальная энтропия может принимать отрицательные значения! Такой вычислительный субъективизм до сего времени не имеет осмысленного физического толкования, а поэтому, в большинстве случаев просто замалчивается. Хотя в [2] предпринята попытка оправдания данного факта тем, что, несмотря на возможную отрицательность дифференциальной энтропии, сумма или разность двух определений энтропий всегда положительна. Однако это оправдание не предотвращает коллапса, который будет рассмотрен ниже при аналитическом определении пропускной способности через среднюю взаимную информацию (соотношение дифференциальных энтропий).

В непрерывном канале входные сигналы источника $x(t)$ являются непрерывными функциями времени, а выходные – $y(t) = x(t) + \xi(t)$ их искаженными суммированием с помехой реализациями. Помеха $\xi(t)$ – также непрерывная функция времени. Пропускная способность непрерывного канала определена в [2], как максимум (по всем возможным распределениям входного сигнала) функции, похожей по своей сути на выражение (1):

$$C = \frac{1}{T} \left(2FT \cdot \max_{f(x)} \{H(Y) - H(Y|X)\} \right), \quad (10)$$

где F – полоса частот, которой ограничен канал; T – продолжительность наблюдения выхода канала; $2FT$ – число степеней свободы, определяемое на длительности T как количество независимых измерений функции с ограниченным спектром по теореме об отсчетах [1, 2]. В формуле (10) $H(Y)$ обозначает энтропию выхода канала, а условная энтропия $H(Y|X)$ определяется выражением (7). Разность, максимум которой ищется в (10), принято назвать средней взаимной информацией между входом и выходом на одно измерение (использование канала):

$$I(X, Y) = H(Y) - H(Y|X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log \frac{f(x, y)}{f(x)f(y)}. \quad (11)$$

Тогда для одного использования канала:

$$C = \max_{f(x)} \{I(X, Y)\}. \quad (12)$$

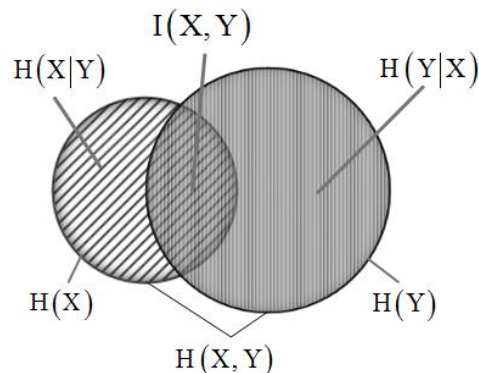


Рис. 1. Взаимосвязь определений энтропии для непрерывного канала

Взаимосвязь рассмотренных информационных определений Шеннона для непрерывного канала удобно рассматривать при помощи диаграммы Венна, представленной на рис. 1. Таким образом, пропускная способность непрерывного канала с аддитивным, статистически не связанным с сигналом шумом на одно измерение равна максимуму средней взаимной информации по всем вариантам распределений входного сигнала. В [2, 4, 7, 8] утверждается, что

$$C = 2F \cdot \max_{f(x)} \{H(Y) - H(\xi)\}, \quad (13)$$

где $H(\xi)$ – энтропия шума. Соответствующая Теорема 16 в [2] постулирует, что при независимых, аддитивно взаимодействующих сигнале и шуме скорость передачи информации на одно использование канала равна разности между энтропией выхода канала и энтропией шума:

$$R = H(Y) - H(\xi), \quad (14)$$

соответственно

$$C = \max_{f(x)} \{R\}. \quad (15)$$

Формулы (10), (12), (13) и (15) представляют собой фактически различные способы определения одной и той же физической величины для различных видов измерений (полного или на одно использование канала). Продолжим рассуждения для гауссова канала в соответствии с логикой изложения из [2], традиционно используемой в учебниках и монографиях по теории информации:

$$H(\xi) = \log \sqrt{2\pi eN}, \quad (16)$$

где N – мощность шума. Для максимизации скорости на основании свойства (9) необходимо потребовать, чтобы распределение источника также было гауссовым с мощностью, равной S . Тогда энтропия источника:

$$H(X) = \log \sqrt{2\pi eS}. \quad (17)$$

Поскольку сигнал и шум, статистически не связаны, то, в силу устойчивости нормального распределения к композиции любого числа суммируемых случайных величин [9], распределение их суммы также будет нормальным с общей мощностью $(S + N)$, что позволяет определить энтропию выхода канала в виде

$$H(Y) = \log \sqrt{2\pi e(S + N)}. \quad (18)$$

В итоге приходим к известной формуле

$$C = F \left[\log(2\pi e(S + N)) - \log(2\pi eN) \right] \quad (19)$$

или

$$C = F \log \left(\frac{S + N}{N} \right). \quad (20)$$

Заметим, что распределение выхода канала будет нормальным в одном единственном случае, если и сигнал, и шум – гауссовы. Поскольку при выводе формулы (20) использованы только функции плотности вероятностей сигнала и помехи (при вычислении дифференциальной энтропии), а о методах обработки при осуществлении приема речи не ведется, то за данной формулой закрепилось определение «пропускная способность гауссова канала» [2 – 8].

Теперь сосредоточимся на странностях поведения компонент аналитического энтропийного определения пропускной способности в форме записи (19). Для этого напомним, что уменьшаемое представляет собой энтропию выхода канала $H(Y)$, а вычитаемое – энтропию шума $H(\xi)$. Что происходит с величиной C при уменьшении мощности шума? Из (20) сле-

дует, что если $F > 0$, то $\lim_{N \rightarrow 0} C = \infty$. При этом формула (19) показывает, что пропускная спо-

собность неограниченно возрастает не за счет роста энтропии выхода канала (которая, напротив, уменьшается), а за счет того, что энтропия шума (вычитаемое в (19)) стремится к минус бесконечности:

$$\lim_{N \rightarrow 0} H(\xi) = -\infty. \quad (21)$$

Это наблюдение противоречит физическому смыслу, заложенному в определении разности (14). При этом смена знака вычитаемого и прибавление к энтропии выхода происходят уже при «слабой» помехе: $N \leq (2\pi e)^{-1}$. Осознать физический смысл этого поведения трудно. Хотя в виде (20) формула пропускной способности демонстрирует монотонность функции $C(N)$ при $N \rightarrow 0$, что позволяет попытаться объяснить рассмотренное явление отличием определения дифференциальной энтропии от дискретной, отмеченным ранее. Однако отсутствие четкого физического толкования этого явления вызывает сомнения по поводу корректности аналитического вывода формулы пропускной способности с использованием понятий дифференциальной энтропии и средней взаимной информации. Как будет ясно в дальнейшем, еще большие сомнения вызывает приписывание этой формуле свойства определять верхнюю границу скоростей передачи информации в гауссовом канале.

2. Геометрическое определение пропускной способности

Спустя год после опубликования «Математической теории связи» К. Шеннон публикует работу [6], в которой дает другой способ определения пропускной способности на основе многомерного геометрического построения пространства сигналов и помех, представленного в «плоском» приближении на рис. 2. Любая реализация непрерывного случайного сигнала длительности T , спектр которого ограничен частотой F , представляется точкой в $n = 2FT$ -мерном пространстве. В «хорошей» системе передачи эти точки – S_i – распределены равномерно внутри гиперсферы с радиусом, определяемым средней мощностью сигнала и размерностью пространства

$$r_S \approx \sqrt{nS} \quad (22)$$

и объемом

$$V_S \approx \frac{\sqrt{\pi}^n}{\Gamma(n/2 + 1)} (\sqrt{nS})^n, \quad (23)$$

где $\Gamma(n/2 + 1)$ – гамма-функция. Для равномерного распределения сигнальных точек можно использовать произвольный выбор n координат – случайных величин с нулевым средним и дисперсией, равной S . При неограниченном возрастании размерности пространства n распределение точек будет монотонно приближаться к равномерному. Это асимптотическое свойство равномерности является основой для построения случайных кодов, практически любой из которых является «хорошим» [7]. Случайная реализация сигнала представляет собой каналный вид кодового слова случайного кода и может быть получена одним из двух способов:

$$S(t) = \sum_{i=0}^{2FT-1} s_i \frac{\sin(2\pi F(t - i \cdot \Delta t))}{2\pi F(t - i \cdot \Delta t)}, \quad \Delta t = 1/(2F); \quad (24)$$

$$S(t) = \sum_{i=1}^{2FT} \left\{ s_{2(i-1)} \sin\left(2\pi F \frac{i}{T}\right) + s_{2(i-1)+1} \cos\left(2\pi F \frac{i}{T}\right) \right\}. \quad (25)$$

Формула (24) представляет собой разложение случайной реализации в базисе отсчетных sinc-функций и обладает непрерывным, ограниченным (эффективно) частотой F спектром. Для случая (25) используется разложение Фурье в ортогональном на интервале T гармонич-

ческом базисе, при этом реализация $S(t)$ периодична на T и, если бесконечно продолжить ее повторение, она будет обладать дискретным спектром, ограниченным сверху величиной F , ненулевые измерения которого расположены с шагом $1/T$ по частоте.

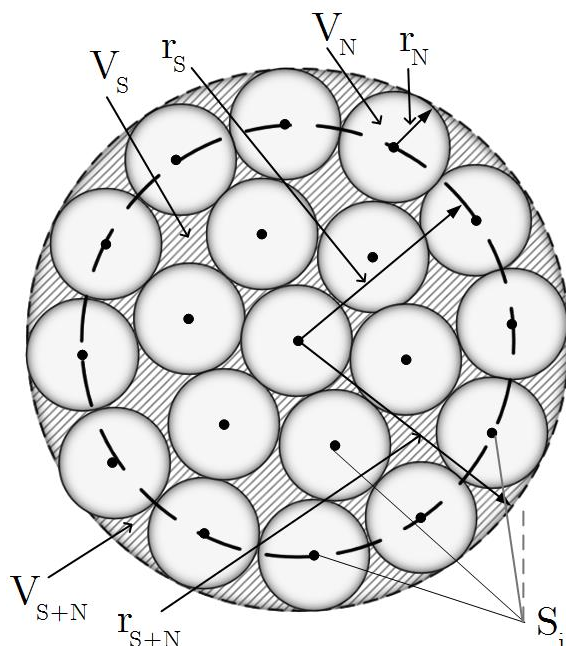


Рис. 2. Геометрическое представление пространства системы передачи информации

Оба способа (24) и (25) могут быть использованы в описании процесса достижения пропускной способности при помощи кодирования, данного Шенноном в [2] (далее цитата № 1): «...Пусть созданы $m = 2^k$ выборок шума, каждая длительности T . Им приписываются двоичные числа от 0 до $m-1$. В передатчике последовательности сообщений разбиваются на группы по k двоичных знаков и для каждой группы в качестве сигнала передается соответствующая выборка шума. На приемном конце эти m выборок известны, и действительно принятый сигнал (искаженный шумом) сравнивается с каждой из них. Выборка, которая имеет наименьшее среднеквадратичное отклонение от принятого сигнала, принимается за переданный сигнал, по которому восстанавливается соответствующее двоичное число. Этот прием эквивалентен выбору наиболее вероятного (апостериори) сигнала...».

Формулы (24), (25) в совокупности с приведенной цитатой представляют собой описание процесса построения и декодирования случайного кода, причем декодирование осуществляется по правилу, которое в настоящее время традиционно называется «*правилом максимального правдоподобия*» (ПМП). При неограниченном возрастании длины блока кода (синхронном увеличении параметров k и $n = 2TF$), если помеха не слишком велика, вероятность ошибки в определении принятого кодового слова может быть сделана сколь угодно малой. Таким образом, в геометрическом определении пропускная способность – это *наибольшая достижимая скорость случайного кода при его декодировании по ПМП* и обеспечении сколь угодно низкой ненадежности.

В геометрической интерпретации лучшего кода (рис. 2) на выходе канала точки S_i , $i \in [0, m-1]$, соответствующие переданным кодовым словам, под воздействием гауссовой помехи смещаются в пределах сфер неопределенности с радиусом

$$r_N \approx \sqrt{nN} \quad (26)$$

и объемом

$$V_N \approx \frac{\sqrt{\pi}^n}{\Gamma(n/2+1)} (\sqrt{nN})^n. \quad (27)$$

В соответствии с проявлением закона больших чисел при увеличении n вероятность нахождения смещенных точек вне сферы с радиусом $r_N + \varepsilon/\sqrt{n}$ стремится к нулю (ε – произвольная малая величина). Сферы неопределенности приобретают более очерченный вид. Шеннон сравнивал их с правильными бильярдными шарами [2, 6, 8]. Поскольку сигналы кодовых слов и помеха не зависят друг от друга, суммарный радиус гиперсферического пространства, вмещающий m сфер неопределенности, характеризуется радиусом и объемом:

$$r_{S+N} \approx \sqrt{n(S+N)}, \quad (28)$$

$$V_{S+N} \approx \frac{\sqrt{\pi}^n}{\Gamma(n/2+1)} \left(\sqrt{n(S+N)} \right)^n. \quad (29)$$

При $n \rightarrow \infty$ можно определить предельное количество непересекающихся сфер, которое можно упаковать в объеме V_{S+N} , причем таким образом, чтобы между ними, практически, не осталось пустых мест:

$$m_C = V_{S+N}/V_N = \sqrt{\frac{S+N}{N}}^n = \sqrt{\frac{S+N}{N}}^{2FT}. \quad (30)$$

Напомним, что, в соответствии с построением кодовых слов по правилам (24) или (25), $2FT = n$ – размерность геометрического пространства кода. Логарифмирование (30) и усреднение за время T дает *предельно достижимую скорость кода*, или (как все-таки утверждает современная теория информации) пропускную способность канала:

$$C = \frac{1}{T} \log m_C = F \log \left(\frac{S+N}{N} \right). \quad (31)$$

Результаты (20) и (31) совпадают по внешнему виду, что будто бы подтверждает определение C как максимально достижимой скорости передачи информации в канале с аддитивным шумом при сколь угодно малой ненадежности. Однако следует отметить, что в логике вывода выражения (31) (как, впрочем, и в сущности рассмотренной выше цитаты) величина C – это предельная скорость лучшего кода, *использующего при декодировании правило максимального правдоподобия*. Если бы это было не так и приемнику не надо было бы хранить в памяти образцы отрезков реализаций сигнала для их использования в сравнениях ПМП, как описано выше в цитате из работы [2], то при $N > S$ достаточно было бы переключиться на прием помехи (какая разница, какой из двух процессов их смеси мы сумеем надежно выделить?), чтобы затем компенсировать ее в выходной смеси канала. Можно обратиться к работе [8] или другим работам, изучающим физический и математический смысл пропускной способности, и обнаружить, что величина C в доказываемых авторами теоремах является именно верхней границей скоростей передачи для *кодов в гауссовом канале, использующих ПМП, но не для гауссова канала при любом способе передачи и обработки сигналов*. В господствующих взглядах теории информации между этими двумя понятиями не делается разницы, поскольку в схеме СПИ, введенной Шенноном [2], присутствие кодера и декодера канала – обязательно по умолчанию. Возможность построения эффективной СПИ, не использующей кодирования, не допускается в принципе! Это резко противоречит практическому наблюдению, отмеченному во введении, о том, что помехоустойчивые коды в режиме исправления ошибок в настоящее время почти не применяются в системах, где ошибки не допустимы. Для уверенного ответа на вопрос: *что же обозначает величина C – просто предельную скорость передачи информации по каналу с аддитивным шумом или предельную скорость передачи информации по каналу при использовании кодирования и правила декодирования ПМП* (и то, и другое – при условии сколь угодно низкой ненадежности), обратимся к анализу математической и логической корректности рассуждений, проведенных при выводе формул (20) и (31). Для объективного анализа необходимо изменить условия, для которых получены формулы (20) и (31), т.е. рассмотреть модели каналов, отличные от гауссовых.

3. Сопоставление аналитического и геометрического определений пропускной способности для негауссова канала

Рассмотрим следующую модель непрерывного канала с полосой частот, ограниченной значением F , и аддитивным, стационарным и независимым от сигнала шумом. Пусть сигнал – гауссов процесс с функцией плотности распределения вероятностей:

$$f_1(x) = \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{x^2}{2S}\right), \quad (32)$$

обладающей математическим ожиданием и дисперсией

$$M[x] = 0, \quad D[x] = S. \quad (33)$$

Энтропия сигнала определяется величиной (17). Помеха в канале такова, что к любому измерению сигнала добавляет случайную ошибку, имеющую равномерную плотность распределения вероятностей в диапазоне $\left[-\frac{a}{2}, \frac{a}{2}\right]$, $a > 0$:

$$f_2(y) = \begin{cases} 1/a, & \text{при } y \in [-a/2, a/2]; \\ 0, & \text{при } |y| > a/2. \end{cases} \quad (34)$$

Соответствующие числовые характеристики распределения (34) составляют:

$$M(y) = 0, \quad D[y] = N = a^2/12. \quad (35)$$

Энтропия такой помехи определяется величиной [3]:

$$H(N) = \log a. \quad (36)$$

Подобной моделью помехи в некоторых случаях может быть описано воздействие квантователя по уровню сигнала при измерении его значений с интервалом дискретизации $\Delta t = 1/2F$ и конечной (больше нуля) величиной a – шага квантования [3].

Теоремой 18 в [2] Шеннон определяет границы величины пропускной способности для произвольного негауссова канала в следующем виде:

$$F \log \frac{S + N_1}{N_1} \leq C \leq F \log \frac{S + N}{N_1}, \quad (37)$$

где N_1 – энтропийная мощность, т.е. мощность эквивалентного гауссова шума, обладающего одинаковой с оригинальным негауссовым шумом энтропией. Для рассматриваемой модели можно вычислить энтропийную мощность, приравнявая между собой величины (16) и (36):

$$N_1 = a^2 / (2\pi e) = \frac{12}{2\pi e} N. \quad (38)$$

Попытаемся теперь вычислить пропускную способность рассматриваемого канала, используя аналитический подход (11) – (14). Энтропия выхода канала в данном случае является дифференциальной энтропией процесса, полученного суммированием двух независимых процессов:

- нормального (сигнал) – с математическим ожиданием и дисперсией (33);
- равномерного (помеха) – с математическим ожиданием и дисперсией (35).

Для вычисления энтропии выхода канала $H(Y)$ необходимо определить функцию плотности распределения вероятностей суммарного процесса $f(y)$, которая в данном случае будет композицией двух распределений [9]:

$$f(z) = \int_{-\infty}^{\infty} f_1(w) f_2(z-w) dw. \quad (39)$$

Использование (32) и (34) в (39) позволяет записать:

$$f(z) = \int_{-a/2}^{a/2} \frac{1}{a} \left[(2\pi S)^{-1/2} \exp\left(-\frac{(z-w)^2}{2 \cdot S}\right) \right] dw = \frac{1}{2a} \left[\operatorname{erf}\left(\frac{a+2z}{\sqrt{8 \cdot S}}\right) + \operatorname{erf}\left(\frac{a-2z}{\sqrt{8 \cdot S}}\right) \right], \quad (40)$$

где $\operatorname{erf}(A) = \frac{2}{\sqrt{\pi}} \int_0^A e^{-t^2} dt$.

В силу независимости двух суммируемых процессов числовые характеристики композиции (40)

$$M[z] = M[x] + M[y] = 0; \quad D[z] = D[x] + D[y] = S + N. \quad (41)$$

Распределение (40) не является гауссовым, хотя очень похоже на него по внешнему виду. Для сравнения на рис. 3 показаны функция плотности распределения вероятностей (ПРВ) (40) и аналогичная гауссова функция для равномошного центрированного процесса при нормированных дисперсиях $a = 2\sqrt{3}$; $S = N = 1$.

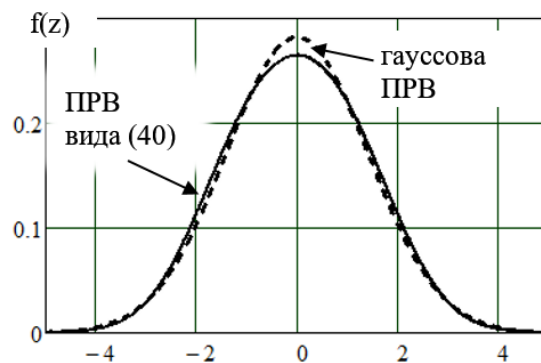


Рис. 3. Сравнение гауссовой и композиционной ПРВ

Естественно, что очень близкими являются и значения дифференциальной энтропии, вычисленные для композиции ПРВ двух нормальных процессов (формула (18)) и для композиции ПРВ рассматриваемого случая. Например, при значениях числовых характеристик, принятых на рис. 3, из (18) имеем:

$$H(Y) = \log \sqrt{2\pi e \cdot 2} = 2,547.$$

Вычисление энтропии распределения (40) дает:

$$H'(Y) = \int_{-\infty}^{\infty} f(z) \log f(z) dz = 2,544,$$

Это означает, что энтропия выхода канала с равномерным шумом, практически, совпадает с аналогичной энтропией гауссова канала, оставаясь все-таки немного меньше

$$H(Y) \approx H'(Y). \quad (42)$$

Данный результат является естественным следствием центральной предельной теоремы теории вероятностей [9]. Можем записать выражение для аналитического вычисления пропускной способности на одно использование рассматриваемого канала, с аддитивной помехой, обладающей равномерной ПРВ, в виде

$$C' = H'(Y) - \log \sqrt{12 \cdot N}, \quad (43)$$

где величина N определяется формулой (35).

Сравнение величины (43) с пропускной способностью на одно использование гауссова канала, определяемой из (19) в одинаковых энергетических условиях, дает

$$\frac{C'}{C} = \frac{H'(Y) - \log \sqrt{12 \cdot N}}{\log \sqrt{2\pi e(S+N)} - \log \sqrt{2\pi eN}}. \quad (44)$$

Пример. Для случая равномошного сигнала и шума $S = N = 1$, рассмотренного на рис. 3, имеем:

- энтропийную мощность, определяемую (38), $N_1 \approx 0,703$;
- границы (37), определенные Шенноном, $0,638 \leq C' \leq 0,755$;
- фактическое значение, вычисленное из (43), $C' \approx 0,751$;
- пропускную способность гауссова канала, определяемую выражением (19) в эквивалентных энергетических условиях, $C = 0,5$;
- отношение пропускных способностей, определяемое формулой (44), $C'/C \approx 1,502$.

Вывод из рассмотренного примера: результаты аналитического энтропийного определения пропускной способности канала с равномерно распределенной помехой приводят к следующему утверждению

пропускная способность канала с равномерно распределенной помехой **в полтора раза превышает** пропускную способность гауссова канала, вычисленную для равномошного сигнала и помехи! (45)

Теперь используем геометрический способ, рассмотренный в п. 2, для определения пропускной способности канала с равномерно распределенной помехой. Для этого сопоставим геометрическое представление и характеристики сфер неопределенности (рис. 2), в пределах которых смещаются сигнальные точки при действии нормальной и равномерной помех. Введем понятие нормированного (на размерность пространства сигналов n) смещения сигнальной точки под действием помех:

– для гауссовой помехи
$$r_n = \left(\frac{1}{n} \sum_{i=1}^n \xi_{n_i}^2 \right)^{1/2}; \quad (46)$$

– для помехи с равномерным распределением
$$r_u = \left(\frac{1}{n} \sum_{i=1}^n \xi_{u_i}^2 \right)^{1/2}; \quad (47)$$

где $\xi_{n_i}, \xi_{u_i}, i \in [1, n]$ – случайное значение i -й координаты аддитивной помехи для нормального и равномерного шума соответственно. Плотности распределения вероятностей этих величин определяются формулами

$$f(\xi_n) = \frac{1}{\sqrt{2\pi N}} \exp\left(-\frac{\xi_n^2}{2N}\right); \quad (48)$$

$$f(\xi_u) = \begin{cases} 1/\sqrt{12 \cdot N}, & \text{при } \xi_u \in [-\sqrt{3 \cdot N}, \sqrt{3 \cdot N}]; \\ 0, & \text{при } |\xi_n| > \sqrt{3 \cdot N}. \end{cases} \quad (49)$$

Нормированные радиусы сфер неопределенности \bar{r}_n и \bar{r}_u для двух рассматриваемых распределений помехи определяются математическим ожиданием случайных величин (46) и (47), являющихся функциями от случайных слагаемых, обладающих ПРВ (48) и (49), а «степень очерченности» сфер – их дисперсией $D[r_n]$ и $D[r_u]$. Для нормальной помехи известен аналитический результат [9]:

$$\bar{r}_n = M[r_n] = \sqrt{\frac{2N}{n}} \left[\Gamma\left(\frac{n+1}{2}\right) / \Gamma\left(\frac{n}{2}\right) \right]; \quad (50)$$

$$D[r_n] = N \left\{ 1 - \frac{2}{n} \left[\Gamma\left(\frac{n+1}{2}\right) / \Gamma\left(\frac{n}{2}\right) \right]^2 \right\}. \quad (51)$$

Аналитическое вычисление аналогичных числовых характеристик для помехи с равномерным распределением \bar{r}_u и $D[r_u]$ сталкивается с трудностями, поскольку многомерная композиционная ПРВ случайной величины (47) является разрывной (кусочно-линейной) функцией. Поэтому эти характеристики были вычислены при помощи статистической модели. Результаты аналитического и статистического исследования характеристик сфер неопределенности для нормального и равномерного распределения координат помехи иллюстрируются на рис. 4 – 6. Результаты вполне ожидаемы, так как следуют из закона больших чисел. На рис. 4 показаны «виртуальные» сечения плоскостью многомерной картины распределения точек смещения для нормального (слева) и равномерного (справа) распределения координат помехи, вычисленные для трех различных значений размерности пространства СПИ при числе испытаний, равном 10^6 .

Изображения сечений сфер получены при одинаковой единичной мощности помехи и, для удобства визуального сопоставления, нормированы (масштабированы) на размерность пространства. Главными выводами из результатов статистического эксперимента и анализа иллюстраций являются следующие:

1) сферы нормальной и равномерной помех обладают примерно равными средними радиусами, при этом величина \bar{r}_u стремится к предельному нормированному значению немного быстрее, чем \bar{r}_n . Это явление иллюстрируется на рис. 5;

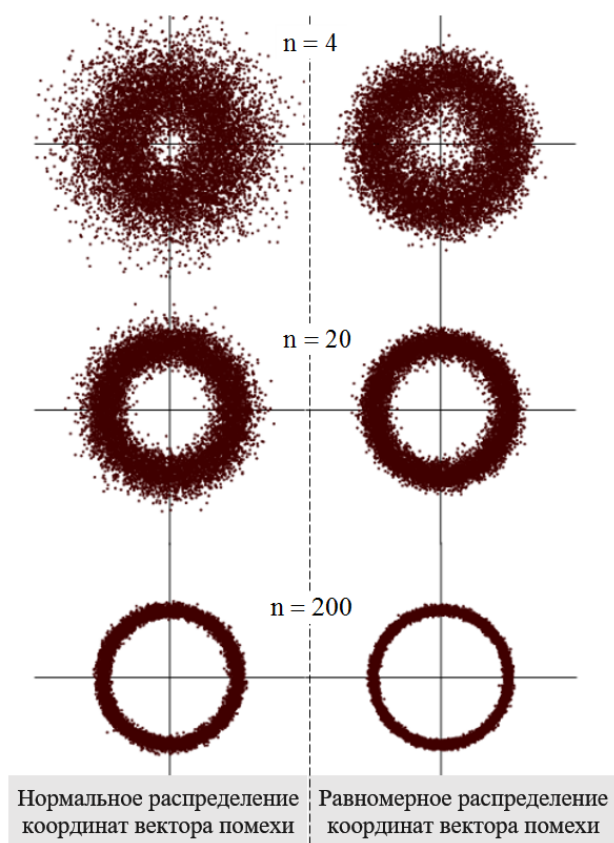


Рис. 4. Проекция нормированных распределений векторов гауссовой и равномерной помех на плоскость

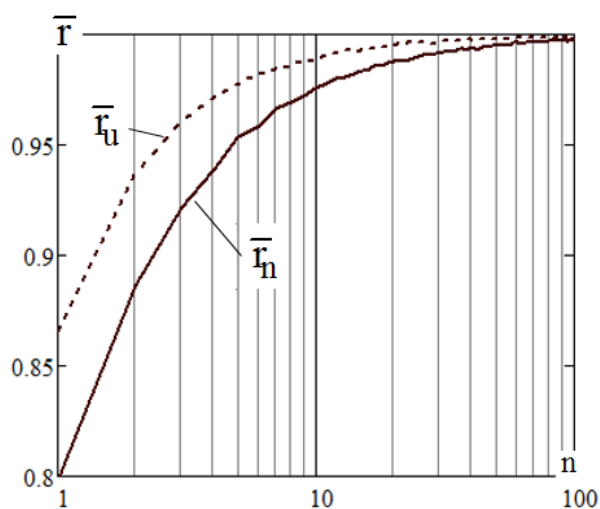


Рис. 5. Зависимости функций \bar{r}_u и \bar{r}_n от размерности пространства n для нормальной и равномерной помех

2) дисперсия разброса значений радиуса сферы равномерной помехи меньше дисперсии сферы нормальной помехи (очерченность контура сфер правой части на рис. 4 более резкая), причем вычисления дают следующее предельное отношение:

$$\lim_{n \rightarrow \infty} (D[r_n]/D[r_u]) = 2,5, \quad (52)$$

т.е. эффективная ширина «кольца» разброса при равномерно распределенной помехе меньше, в среднем, в $\sqrt{2,5}$ раз (практически, при любой размерности пространства n), чем аналогичный параметр при нормальной помехе. Предельные абсолютные значения дисперсий радиусов сфер составляют:

$$\lim_{n \rightarrow \infty} \{n \cdot D[r_n]\} = N/2, \quad \lim_{n \rightarrow \infty} \{n \cdot D[r_u]\} = N/5. \quad (53)$$

Для $N=1$ это явление иллюстрируется графиками на рис. 6. Для нормально распределенной помехи абсолютная дисперсия радиуса с ростом n незначительно возрастает, стремясь к предельному значению «снизу», а для равномерно распределенной помехи – убывает, стремясь к пределу (53) «сверху».

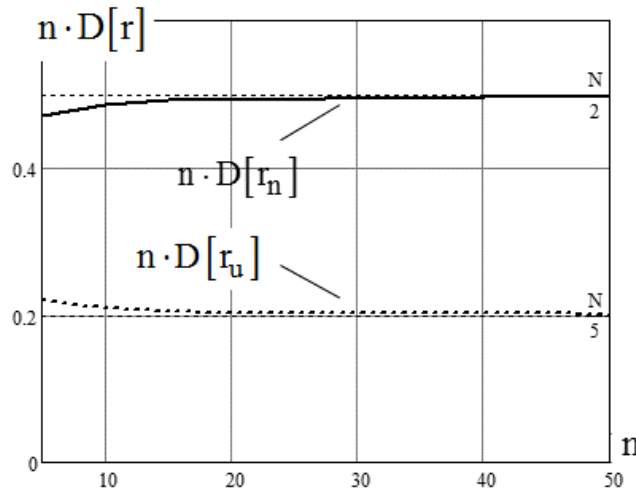


Рис. 6. Дисперсия радиусов сфер неопределенности для нормальной и равномерной помехи как функция размерности пространства

Наконец, главный очевидный вывод:

3) средние радиусы сфер неопределенности для рассматриваемых типов ПРВ помех асимптотически совпадают:

$$\lim_{n \rightarrow \infty} \bar{r}_n = \lim_{n \rightarrow \infty} \bar{r}_u = \sqrt{n \cdot N}. \quad (54)$$

Данный результат – следствие закона больших чисел. Он, естественно, может быть обобщен **для любых видов** центрированных ПРВ сигнала и помехи, т.е. для любых непрерывных каналов с аддитивной, статистически не связанной с сигналом помехой. На параметры геометрического представления СПИ при $n \rightarrow \infty$ влияют только значения средних мощностей непрерывного сигнала и помехи, но не вид их распределения!

Поскольку по аналогичным причинам радиус гиперсферы пространства на выходе негауссова канала также совпадает с величиной, определяемой выражением (29), то используя (27), (29) и (30) приходим к той же самой величине пропускной способности канала с равномерно распределенной помехой: $C' = C = F \log \frac{S+N}{N}$, что противоречит определению (43) и выводу (45). Таким образом, две работы К. Шеннона [2] и [6], опубликованные с интервалом в один год, в применении к негауссову каналу – противоречат друг другу. На вопрос «какой из двух способов определения пропускной способности: аналитический (энтропийный) или геометрический является корректным?» – существует однозначный ответ: конечно геомет-

рический. Корректность геометрического подхода легко проверяется методом статистического моделирования случайного кода [12]. Аналитический способ дает результат, совпадающий с результатом геометрического способа в одном единственном случае, когда сигнал и помеха – гауссовы процессы. Это всего лишь совпадение, природа которого объясняется свойствами нормального распределения, имеющего особое значение в теории вероятностей и случайных процессов. В силу отмеченных причин методология использования «энтропийной мощности» и границы, определяемые (37) являются не корректными.

Результаты анализа гауссова и негауссова канала (как, впрочем, и любой другой модели) показали, что эти каналы имеют одинаковое значение пропускной способности $C = C'$, величина которой зависит только от отношения сигнал/шум (по мощности) и полосы частот канала. Поэтому приписывание свойства определения величиной C (20) предельной скорости передачи информации именно в *гауссовом* канале с аддитивным шумом, по меньшей мере, некорректно.

Истинный физический смысл пропускной способности в геометрическом выводе заключается в определении *предельной скорости передачи информации по каналу с любым видом аддитивного шума при использовании канального кодирования и правила максимального правдоподобия при декодировании*. Следовательно, пропускная способность – это не характеристика канала, а *естественный предел, возникающий для любой модели непрерывного канала, как только мы решим использовать кодирование информации при передаче* (в смысле принятия решения по результатам сопоставления выхода канала с известными образцами допустимых реализаций сигналов).

При этом, как следствие, возникает необходимость разбиения пространства сигналов на выходе канала на области «похожести», которыми, собственно, и являются сферы неопределенности в геометрическом представлении на рис. 2. Эти области не будут пересекаться до тех пор, пока при фиксированном бюджете мощности передатчика мощность помехи не будет превышать некоторого допустимого значения. Данное допустимое значение, фактически, и определяет так называемую пропускную способность (на самом деле – предельную скорость лучшего достижимого кода). Источником фундаментальных ограничений в существующей парадигме теории информации является господствующая аксиоматическая неизбежность применения кодов и осуществление решений о приеме по принципу «наибольшей похожести». Если сформулировать этот вывод наиболее лаконичным способом, то весьма скромные современные достижения теории передачи информации – следствие безальтернативного применения, так называемого, *правила максимального правдоподобия*.

В заключении текущего раздела приведем соображения, которые являются дополнительным аргументом доказательства, в определенной степени, некорректности существующего аналитического вывода пропускной способности, как максимума средней взаимной информации, рассмотренного в п. 1. В тексте цитаты из [2], приведенной выше в п. 2, кодирование рассматривается как процесс сопоставления выборки шума с одной из $M = 2^k$ комбинаций символов источника. При этом, очевидно, правильным будет определение энтропии этой выборки вовсе не формулой (17), а как неопределенности дискретного выбора (согласно принципу (2)), т.е.

$$H(X) = \log 2^k = k. \quad (55)$$

Тогда именно это определение должно использоваться в вычислениях (17) – (20), что приводит к очередному коллапсу, поскольку в одном выражении будут присутствовать определения энтропии для дискретного и непрерывного выбора, которые существуют в различных системах измерений.

4. Правило максимального правдоподобия

Теорема Крамера (1740 г.):

Не существует другого метода обработки результатов эксперимента, который дал бы лучшее приближение к истине, чем метод максимального правдоподобия.

Правило (метод) максимального правдоподобия имеет название, вполне соответствующее его роли в статистическом оценивании реализаций случайного опыта и принятии решений в условиях многоальтернативных гипотез. Современная парадигма передачи информации во всех известных практических приложениях рассматривает принятие решений о состоянии выхода зашумленного канала в условиях равновероятных гипотез, т.е. все сообщения источника предполагаются равновероятными, а действие на них шума в канале – одинаковым (симметричным). Это объясняет безальтернативность ПМП по сравнению с другими статистическими методами и критериями принятия решений. Без особого преувеличения можно утверждать, что правило (критерий) максимального правдоподобия перекочевало в статистическую теорию связи из нашего жизненного опыта. Мы всегда пытаемся расслышать фразу в условиях мешающего шума или различить предмет в условиях плохой видимости, подсознательно включая алгоритм под названием «на что (из известного нам) это больше всего похоже?». Именно этим объясняется аксиоматично объективное использование ПМП во всех стандартных приложениях теории передачи информации без особых забот об аргументации такого выбора.

Цитата из [2], к которой мы уже обращались (см. п. 2 данной работы), отражает вполне оправданную из соображений нашего физиологического опыта позицию К. Шеннона о том, что декодер на выходе канала должен принимать решение о принятом кодовом слове (сигнале) на основании сопоставления близости (в среднеквадратическом смысле) принимаемой выборки случайного процесса на выходе канала с имеющимися у приемника образцами.

Такой же подход наблюдается при описании идеального (по Котельникову) приемника при некодированной модуляции [1] (цитата № 2, обозначения переменных изменены): «... будем считать, что в зависимости от воздействующего на приемник суммарного колебания $y(t)$ он обязательно воспроизведет одно из возможных значений сообщения $S_1(t), \dots, S_m(t)$ Очевидно ... всю область возможных значений $y(t)$ можно разбить на m не перекрывающихся друг друга областей. ... В зависимости от конфигурации областей, которая определяется приемным устройством, будет получаться больше или меньше правильно воспроизведенных сообщений. ... Приемник, характеризуемый такими (прим. авт.: правильно выбранными) областями и дающий вследствие этого минимальное число неправильно воспроизведенных сообщений при наложении помехи, мы будем называть *идеальным*».

Следовательно, образующим постулатом современной теории потенциальной помехоустойчивости [1], как и теории помехоустойчивого кодирования [2], является правило обработки зашумленных сигналов (кодов) на основе максимального правдоподобия (или максимального сходства), которое используется авторами, как фундамент выстраиваемых затем теорий.

Если значения априорных вероятностей сообщений источника являются одинаковыми, то математическая формулировка ПМП при выборе k -й гипотезы из m альтернатив имеет вид:

$$\frac{f(S_k|y)}{f(S_i|y)} > 1, \text{ для всех } i \in [1, m], i \neq k, \quad (56)$$

где $f(S_i|y)$ – функция правдоподобия, записанная для сообщения S_i . Задача отыскания наиболее достоверного решения сводится к максимизации функций правдоподобия, и, в некоторых случаях, может иметь аналитическое (не переборное) решение на основе известных

из математического анализа методов исследования на экстремум. В рассматриваемых приложениях для непрерывного канала (см. выше цитаты 1 и 2) на длительности T функция правдоподобия для сообщения S_i может быть выражена через евклидово (гильбертово) расстояние:

$$f(S_i|y) = \left\{ \int_T [S_i(t) - y(t)]^2 dt \right\}^{-1/2}. \quad (57)$$

В соответствии с принципом максимального правдоподобия (наибольшей схожести) гипотеза, обладающая максимумом функции (57), считается истинной [1, 2]. Прибегая к подобному правилу, мы автоматически вводим предельное ограничение на допустимую интенсивность мешающих помех, т.е. *ограничиваем снизу отношение сигнал/шум при котором еще не будет происходить выход отображающей принимаемый сигнал точки за пределы собственной области схожести*. Данный процесс ограничения порождает все основные утверждения и, так называемые, фундаментальные пределы в теории передачи информации. Эти пределы, главным из которых, несомненно, является пропускная способность, к сожалению, чрезвычайно жестки, что, собственно и является причиной весьма скромных достижений теории передачи информации.

Какова величина вероятности P – сохранения схожести процесса на выходе канала переданному сообщению при малых отношениях сигнал/шум? Ответ очевиден – весьма мала. Пусть каналный алфавит позволяет передавать m различных сообщений (сигналов), которые могут появляться одинаково часто. Тогда, при фиксированной мощности сигнала S , при возрастающей мощности шума N справедливо записать:

$$\lim_{N \rightarrow \infty} P = m^{-1}; \quad \lim_{m \rightarrow \infty} P = 0. \quad (58)$$

При большой помехе (скорости выше пропускной способности) процесс на выходе канала с вероятностью, близкой к единице, не является похожим на истинно переданное сообщение, поскольку с равной вероятностью его отображающая точка может оказаться в области схожести, практически, любого из m возможных сообщений. При упаковках сигнальных точек в n -мерном пространстве, близких к плотнейшим [11], число сфер неопределенности, примыкающих к сфере схожести истинно переданного сигнала, может оказаться слишком большим. Это не позволяет создать многомерные упорядоченные манипуляционные коды (типа кода Грея), минимизирующие количество искажаемых двоичных символов при ошибках трансформации истинного сообщения в ближайшее к нему в пространстве СПИ. Например, при $n = 24$ известна плотнейшая укладка на основе решетки Лича, построенная с использованием двоичного кода Голя [10, 11], в которой к поверхности одной сферы примыкают 196560 окружающих ее сфер. Если построить на основе этой решетки любые избыточные помехоустойчивые $(24, k)$ коды при $k = 1, 2, \dots, 18$, то можно обеспечить попарную равноудаленность почти всех кодовых (сигнальных) точек. Даже при небольшом превышении пропускной способности (небольшом перекрытии сфер неопределенности) прием по ПМП любого кодового слова на выходе канала является почти равновероятным, и, практически, не зависит от переданного слова (сообщения). Применение правила максимального правдоподобия в таких условиях заведомо приведет к ошибке приема. Налицо парадокс и противоречие: с одной стороны, ПМП является лучшим способом приема, минимизирующим вероятность ошибок при слабых помехах; с другой – это правило само является причиной ограничений на допустимую скорость и/или мощность помех. Возможны ли модификации правила принятия решений при использовании кодирования и вероятностной оценки состояния выхода канала?

5. Могут ли коды работать без правила максимального правдоподобия?

Оценку возможности изменения решающего правила, когда истинно переданным считается сообщение, не являющееся наиболее близким к реализации на выходе канала (в среднеквадратическом смысле), удобно провести с представлением пространства сообщений Пуассоновым полем точек [12].

При построении случайного или алгебраически упорядоченного кода его кодовая книга (в виде множества сигнальных точек) образует в n -мерном пространстве случайное (Пуассоново) поле точек так как всегда выполняются следующие условия:

1) при фиксированном среднем энергетическом бюджете передатчика все точки кодовых слов размещаются в ограниченном объеме многомерного пространства, причем это размещение при увеличении n асимптотически приближается к равномерному (для случайного кода), т.е. плотность поля точек постоянна по всему объему кодового пространства;

2) вероятность появления произвольного числа точек в любом объеме пространства не зависит от того, сколько точек попало в любые объемы, не пересекающиеся с данным;

3) вероятность попадания в элементарный объем двух или более точек пренебрежимо мала по сравнению с вероятностью попадания одной точки.

Предположим, что скорость передачи в произвольном гауссовом канале превышает его пропускную способность. В геометрическом представлении это приведет к взаимному пересечению сфер неопределенности, показанному для фрагмента выходного пространства канала на рис. 7. Для моделирования ситуации используем известное [9] аналитическое описание ПРВ $\varphi(\Delta)$ случайной величины помехового смещения $\Delta = \sqrt{n} \cdot r_n$ (здесь r_n определяется формулой (46)):

$$\varphi(\Delta) = \frac{2\Delta^{n-1}}{\Gamma\left(\frac{n}{2}\right)\sqrt{2N^n}} \exp\left(-\frac{\Delta^2}{2N}\right). \quad (59)$$

Числовые характеристики $\varphi(\Delta)$ находятся из (50), (51) и составляют:

$$M[\Delta] = \sqrt{n} \cdot \overline{r_n}, \quad D[\Delta] = n \cdot D[r_n]. \quad (60)$$

Пусть по каналу в условиях действия аддитивной помехи было передано сообщение, соответствующее точке 1. Смещающее действие помехи таково, что приемнику доступна для наблюдения на выходе канала точка 2, при этом допустим, что величина смещения составляет $\Delta = (M[\Delta] + \delta)$.

Очевидно, что в этой ситуации применение ПМП выявит ближайшую к принятой точке 2 – точку 3, как истинно переданную, что приведет к ошибке.

Модифицируем правило принятия решения следующим образом: полагая наблюдаемую на выходе канала точку 2 центром, восстановим вокруг нее поверхность сферы с радиусом $M[\Delta]$. Затем, просматривая поочередно все точки кодовой книги, выявим точку, расположенную ближе остальных к поверхности построенной сферы. Эта точка и будет считаться истинно переданной. В соответствии с описанным правилом, на рис. 7 истинно переданной является точка 1, расположенная на расстоянии δ от поверхности вспомогательной сферы, что соответствует безошибочному решению приемника в данном примере. Назовем данное правило принятия решения «правилом сферы неопределенности» (ПСН). В соответствии с этим правилом истинной считается не сообщение, являющееся наиболее похожим на наблю-

даемую реализацию выхода канала, а сообщение, которое находится ближе остальных к поверхности сферы с радиусом $M[\Delta]$, очерченной вокруг точки наблюдаемого выхода.

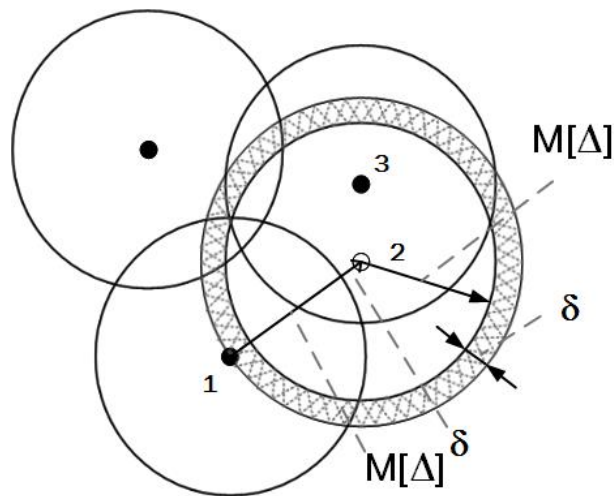


Рис. 7. Геометрическая иллюстрация правила сферы неопределенности (ПСН)

Функция правдоподобия произвольного сигнала S_i для ПСН может быть записана так:

$$f(S_i | y) = \left| \int_T [S_i(t) - y(t)]^2 dt - M[\Delta]^2 \right|^{-1/2}. \quad (61)$$

Истинно переданным считается сигнал, обладающий максимальным значением функции (61). Описанное правило будет гарантированно приводить к безошибочному решению только в случае, если вспомогательная сфера вокруг принятой точки (на рис. 7 – это точка 2) будет иметь радиус, точно совпадающий с величиной помехового смещения переданной точки, т.е. в случае, если абсолютно точно известна мощность помехи, добавившейся к переданному сигналу (кодovому слову) в конкретной наблюдаемой реализации выхода канала. Однако, так как точное знание мощности компоненты помехи в принимаемой конкретной реализации смеси сигнала и помехи невозможно, то вспомогательная сфера может быть очерчена только с радиусом, равным ее математическому ожиданию $M[\Delta]$. Это может привести к ошибочному решению в случае, если в прослойке толщиной δ – между концентрическими сферами с радиусами $\Delta = (M[\Delta] + \delta)$ и $M[\Delta]$ (заштрихованное кольцо на рис. 7) окажется любая другая точка из набора кодовой книги. Вероятность ошибочного решения может быть оценена, как функция от S , N , n на основании свойств поля Пуассона. Ошибка произойдет, если внутри пространства между двумя концентрическими сферами окажется хотя бы одна кодовая точка. Вероятность этого события для рассматриваемого поля точек:

$$P(\lambda, \Delta) = 1 - \exp(-\lambda(m) \cdot V(\Delta)), \quad (62)$$

где λ – плотность поля, содержащего m точек:

$$\lambda(m) = m/V_{S+N}; \quad (63)$$

величина V_{S+N} определяется выражением (29); $V(\Delta)$ – объем концентрической прослойки около вспомогательной сферы:

$$V(\Delta) = \frac{\sqrt{\pi}^n}{\Gamma(\frac{n}{2} + 1)} \begin{cases} \{M[\Delta]^n - \Delta^n\}, & \text{при } 0 \leq \Delta \leq M[\Delta]; \\ \{\Delta^n - M[\Delta]^n\}, & \text{при } \Delta > M[\Delta]. \end{cases} \quad (64)$$

Используя (62), (63) и усредняя результат в соответствии с распределением (59), можно вычислить вероятность ошибки при декодировании по ПСН:

$$P_{er} = \int_0^{\infty} \varphi(\Delta) P(\lambda, \Delta) d\Delta. \quad (65)$$

При заданных значениях S, N и n число сигнальных точек m_C в пространстве кода, соответствующее достижению пропускной способности, определяется из (30), при этом плотность поля точек $\lambda(m_C)$ вычисляется по формуле (63). Вводя коэффициент α – изменения скорости передачи на одно использование канала R , можно смоделировать ситуации, когда скорость превышает пропускную способность C , что приводит к пересечению сфер неопределенности:

$$R > C \rightarrow R = \alpha \cdot C \rightarrow \alpha > 1 \Rightarrow m = (m_C)^\alpha; \quad (66)$$

или, напротив, не достигает пропускной способности (сферы неопределенности с некоторым запасом не пересекаются):

$$R < C \rightarrow R = \alpha \cdot C \rightarrow \alpha < 1 \Rightarrow m = (m_C)^\alpha. \quad (67)$$

В данных выражениях аргументом, регулирующим моделируемую скорость, является число точек m различных сигналов (кодовых слов) при фиксированном объеме сигнального пространства. При $m > m_C$ пропускная способность превышена, а при $m < m_C$ – скорость передачи меньше пропускной способности. Коэффициент α в (66) и (67) расположен в показателе степени, так как скорость передачи измеряется логарифмом от m .

На рис. 8 представлены результаты вычисления вероятности ошибочного решения (65) для правила ПСН при различных значениях коэффициента α и $S = N = 1$.

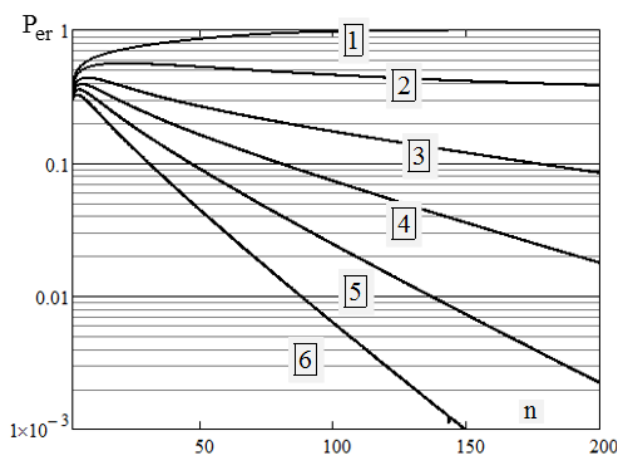


Рис. 8. Зависимости $P_{er}(n)$ при использовании ПСН при различных скоростях: $\boxed{1} \rightarrow R = 1,05 \cdot C$; $\boxed{2} \rightarrow R = 0,95 \cdot C$; $\boxed{3} \rightarrow R = 0,8 \cdot C$; $\boxed{4} \rightarrow R = 0,7 \cdot C$; $\boxed{5} \rightarrow R = 0,6 \cdot C$; $\boxed{6} \rightarrow 0,5 \cdot C$

Увы, основной вывод из анализа кривых на рис. 8 неутешителен – правило ПСН (такое привлекательное в примере, разобранным на рис. 7) приводит к такому же результату, как и правило ПМП! При $R > C$ вероятность ошибки при росте размерности пространства (числа степеней свободы или длины блока случайного кода) монотонно стремится к единице. При $R < C$ – вероятность ошибки может быть сделана сколь угодно малой при соответствующем увеличении n . Данный результат объективно объясняется свойством многомерных сфер концентрировать практически весь свой объем в малой области, прилегающей к поверхности. В этой же области, непосредственно окружающей поверхность вспомогательной сферы,

ПРВ, определяемая выражением (59), достигает своих наибольших значений. Поэтому при увеличении размерности пространства эффективный объем прослойки вокруг вспомогательной сферы растет быстрее, чем падает плотность поля точек. Конечно, можно попытаться сформулировать другие критерии и правила принятия решения, но, очевидно, результат будет не лучше результата ПМП (теорема Крамера). На основании проведенного моделирования можно сделать однозначный вывод – ПМП является лучшим и, возможно, единственно приемлемым правилом статистических решений для кодов – альтернатив ему по-видимому, не существует.

Обсуждение результатов и выводы

Основные выводы из результатов анализа математической и физической сущности пропускной способности, а также противоречий в современной теории передачи информации сконцентрированы в следующей последовательности рассуждений, являющихся логическим обобщением промежуточных выводов.

1. Вероятностно-энтропийный подход к аналитическому определению пропускной способности непрерывных каналов, использующий понятие средней взаимной информации между входом и выходом (5) – (15), можно считать корректным только при условии гауссовости распределений источника и канала (16), (17). Поскольку применение этого подхода для не гауссовых моделей непрерывных каналов приводит к ошибочным результатам (39) – (45), то во многих опубликованных к настоящему времени работах делается неоправданный вывод о практической невозможности аналитического определения пропускной способности для подобных моделей.

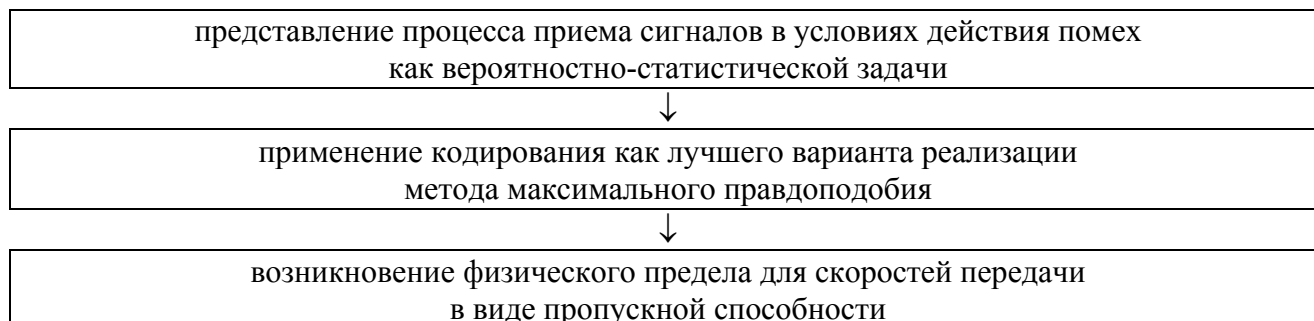
2. Математическое определение (31) корректно описывает величину пропускной способности любого непрерывного канала, шум в котором – стационарный случайный процесс. Вид распределения шума не влияет на величину пропускной способности, которая полностью определяется отношением сигнал/шум и полосой частот. Различные распределения шума проявляются только в изменении скорости приближения к пропускной способности при увеличении длительности отрезков шумовых последовательностей случайного кода.

3. Корректное геометрическое определение пропускной способности описывает ее физическую сущность как предельную скорость передачи информации по каналу с любым видом аддитивного шума при использовании канального кодирования/декодирования и применении правила максимального правдоподобия для декодирования. При этом предполагается, что реализация ПМП осуществляется путем сравнения наблюдаемого на выходе канала сигнала с имеющимися у приемника эталонными образцами. Пропускная способность – физический предел только для систем, использующих метод максимального правдоподобия.

4. Правило максимального правдоподобия является единственным и лучшим правилом принятия решений при декодировании. В то же время пропускная способность – косвенное определение нижней границы отношения сигнал/шум, при котором с большой вероятностью еще не происходит помеховое смещение точек, отображающих сообщения в многомерном пространстве выхода канала, за пределы фиксированных «областей похожести». Существование этих областей определено сущностью метода максимального правдоподобия. Таким образом, правило максимального правдоподобия, с одной стороны, – лучшее правило принятия статистических решений, а с другой – причина появления физического предела в виде пропускной способности. Отказаться от использования ПМП, приводящего к появлению фи-

зического предела скоростей передачи информации, в условиях, когда работа приемника заключается в решении вероятностно-статистической задачи, – невозможно!

5. Приведенные рассуждения порождают следующую логическую причинно-следственную цепочку факторов, приведших к кризису в развитии теории передачи информации:



Как показано в данной работе, разорвать представленную цепочку нельзя:

– невозможно превысить пропускную способность, не отказавшись от метода максимального правдоподобия;

– невозможно отказаться от максимального правдоподобия в условиях, когда работа приемника рассматривается, как решение вероятностной задачи.

Таким образом, в корне рассматриваемых проблем лежит вероятностный подход к организации приема сигналов в условиях действия помех, альтернатив которому обычно не возникает даже в самых смелых предположениях. Однако это не так, просто мы используем не все возможности, представляемые природой для обработки зашумленных цифровых сигналов в непрерывных каналах. Но об этом, как уже анонсировалось во введении планируется в следующих публикациях.

Список литературы: 1. Котельников В.А. Теория потенциальной помехоустойчивости / В. А. Котельников. – М. : Госэнергоиздат, 1956. – 152 с. 2. Shannon C.E. A Mathematical Theory of Communication / C.E. Shannon. – Bell Syst. Tech., July-October, 1948. – Vol. 27. – P. 379-423, 623-656. 3. Помехоустойчивость и эффективность систем передачи информации / А.Г. Зюко, А.И. Фалько, И.П. Панфилов, В.Л. Банкет, П.В. Иващенко ; под ред. А.Г. Зюко. – М. : Радио и связь, 1985. – 272 с. 4. Галлагер Р. Теория информации и надежная связь / Галлагер Р. – [пер. с англ. ; под ред. М.С. Пинскера и Б.С. Цыбакова]. – М. : Сов. радио, 1974. – 720 с. 5. Добрушин, Р.Л. Теория передачи информации / Р.Л. Добрушин, Б.С. Цыбаков // Вестник АН СССР. – 1998, № 6. – С. 76 – 81. 6. Shannon C.E. Communication in the presence of noise / C.E. Shannon // Proc. IRE, vol. 37, January, 1949. – pp. 10–21. 7. Verdu, S. Fifty Years of Shannon Theory / S. Verdu // IEEE Transactions on Information Theory, Vol. 44, № 6, October 1998. – pp. 2057 – 2078. 8. Verdu, S. A General Formula for Channel Capacity / S. Verdu, Te Sun Han // IEEE Transactions on Information Theory, Vol. 40, № 4, July 1994, – pp. 1147 – 1157. 9. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. – М. : Высш. шк., 2000. – 480 с. 10. Хэмминг, Р.В. Теория кодирования и теория информации / Р.В. Хэмминг : пер. с англ. – М. : Радио и связь, 1983. – 176 с. 11. Конвей, Дж. Упаковки шаров, решетки и группы. В 2 т. / Дж. Конвей, Н. Слоэн – М. : Мир, 1990. 12. Рассомахин С. Г. Технология псевдослучайного кодирования в сетевых коммуникационных протоколах канального уровня / С. Г. Рассомахин // Системы обработки информации. – 2012. – Вып. 3(101), т.2. – С. 206 – 211.

Харьковский национальный
университет имени В.Н. Каразина

Поступила в редколлегию 18.01.2017