

# Глава 4. Создание и развитие теории информации

## § 1. Предыстория

Годом рождения теории информации следует считать 1948 г., когда выдающийся американский ученый **Клод Шеннон** опубликовал свою знаменитую статью «Математическая теория связи» об основных закономерностях передачи информационных сообщений по каналам связи.

Подобно тому, как за открытием Исааком Ньютоном законов механики последовало бурное развитие многих современных технологий (например, всех видов наземного и воздушного транспорта, космической техники и т.д.), так и открытие новых закономерностей передачи информационных сообщений создало предпосылки быстрого и глобального развития телекоммуникаций. Эти закономерности связаны с двумя разными аспектами проблем, которые с давних пор стояли перед инженерами, разрабатывающими новые системы связи.

Первая из них касалась возможностей устранения избыточности передаваемых сообщений, что позволило эффективнее использовать каналы связи, вторая — возможностей передачи сообщений с минимальными искажениями по каналам связи с шумом.

Для решения обеих проблем необходим был адекватный аппарат, позволяющий математически описывать сигналы, передаваемые по каналам связи. Ни одно из крупных научных направлений, естественно, не возникает на пустом месте. Фундамент, на котором К. Шеннон построил свою теорию, подготавливался его предшественниками, в работах которых, выполненных в первой трети XX столетия, были сформулированы новые понятия и установлены важные принципы теории связи.

**Математическое описание сигналов, передаваемых по каналу связи.** Отметим две важные работы, которые в начале XX в. формировали представление инженеров о свойствах сигналов, используемых для передачи информации. Дж. Карлсон в 1922 г. ввел понятие спектра сигналов и определил спектр сигналов с ЧМ [1], подведя черту под спором, который инженеры вели несколько лет, о возможности сокращения с помощью ЧМ полосы частот, занимаемой системой передачи сообщения в эфире. Интересно отметить, что на самом деле за шесть лет до Дж. Карлсона важнейшее для теории связи понятие спектра сигналов было введено применительно к сигналам с АМ академиком М.В. Шулейкиным. Однако его работа [2] была опубликована в малоизвестном журнале и не стала доступной широкому кругу специалистов. В 20-е гг. дискуссия об адекватности спектральных представлений в теории связи была весьма жаркой. Так, знаменитый изобретатель вакуумного диода американский инженер Дж. Флеминг резко выступал против их использования и считал возможным, применяя ЧМ, сократить занимаемую системой связи в эфире полосу частот. Важной работой оказалась статья Д. Габора [3], в которой в теорию связи было введено, по-видимому,

впервые плодотворное понятие комплексного аналитического сигнала, действительную часть которого представляет сам сигнал, а мнимая — является его преобразованием Гильберта. Эти понятия сегодня играют в теории связи ключевую роль.

Один из основных законов теории связи был установлен в 1924 г. **Г. Найквистом** (США) [4] и **К. Купфмюллером** (Германии) [5]. Согласно этому закону, по каналу с шириной полосы  $F$  за время  $T$  можно передать не более  $2FT$  значений телеграфного сигнала. Математически этот закон выражает важнейшую закономерность в теории связи следующим образом:

$$N \geq 2FT,$$

где  $N$  — число отсчётов, необходимое для представления временного отрезка сигнала длительностью  $T$ , а  $F$  — полоса частот, занимаемая телеграфным сигналом. При данном критерии информация не утрачивается и этот сигнал может быть полностью восстановлен. Установленные Найквистом основополагающие принципы теории связи оказали огромное влияние на прогресс в области телекоммуникаций, связанный с переходом от аналоговых систем передачи сообщений (аналоговые сигналы вырабатываются большинством источников информации — микрофонами, телевизионными камерами и т.п.) к цифровым. В цифровых системах каждый из отсчётов сигнала преобразуется в цифровую форму и в виде последовательности символов (чаще всего бинарных) передаётся по каналу связи.

К знаковым работам, относящимся к предистории теории информации, относится работа академика **В.А. Котельникова** [6], который в 1933 г. установил знаменитую теорему, позволяющую представить любую функцию  $S(t)$  с ограниченной полосой частот  $F$  своими отсчётами, взятыми через интервалы времени  $\tau = 1/2F$  следующим образом:

$$S(t) = \sum_{n=-\infty}^{\infty} S(n\tau) [\sin 2\pi F(t - n\tau) / 2\pi F(t - n\tau)].$$

Позже, в 1948 г., эта же формула независимо была установлена **Клодом Шенноном**. В честь этих крупнейших учёных XX в. данную формулу иногда называют формулой Котельникова—Шеннона—Найквиста. Важную роль в развитии теории связи сыграли исследования **Д.В. Агеева** [7], который рассмотрел возможности передачи в одном канале нескольких сообщений, осуществляющих модуляцию различных ортогональных сигналов. Он был, по-видимому, первым, кто плодотворно стал использовать в теории связи понятия многомерного пространства. Позже эти понятия эффективно применяли в своих основополагающих исследованиях **В.А. Котельников** и **К. Шеннон**.

**Сокращение избыточности сообщений** привлекало внимание людей еще в древние времена. Так, со времен Юлия Цезаря для ускорения записи его речей использовалась стенография — способ кодирования слов речи, при котором для записи наиболее часто употребляемых оборотов речи применялись наиболее простые и короткие символы.

Для сокращения избыточности телеграфных сообщений знаменитый американский изобретатель телеграфа **Самуэль Морзе** создал в 1837 г. код, который в его честь называют кодом Морзе. Этот код был построен таким образом, что наиболее часто встречающимся в тексте буквам английского алфавита были присвоены наиболее

короткие последовательности, состоящие из точек и тире. Он позволял передавать телеграфные сообщения по каналу связи, используя наименьшее в среднем число символов, т.е. затрачивая на передачу меньшее время. Этот код, учитывающий распределение вероятностей в текстах букв английского алфавита, явился первым практическим примером статистического кодирования дискретного источника сообщений. О высокой эффективности кода Морзе говорит то, что он до сих пор находит практическое применение.

Речевые сигналы также обладают большой избыточностью. В 1939 г. американский инженер Г. Дадли (США) изобрел вокодер — устройство, использующее особенности речи для сокращения пропускной способности канала связи, необходимой для ее передачи, в 5–10 раз [8].

Хотя отдельные проблемы устранения избыточности в сообщениях были решены задолго до появления теории информации, это нисколько не умаляет значения открытых Шенноном общих принципов сжатия данных, относящихся к источникам информации произвольной природы. Это направление теории информации имеет важное прикладное значение. В последние годы на основе этих принципов разработан ряд весьма эффективных практических методов для сжатия сигналов телеметрии, фототелеграфии, телефонии, телевидения, звукового вещания и пр. Следует подчеркнуть, что сжатие данных чрезвычайно важно не только для электросвязи, где оно позволяет эффективнее использовать каналы связи, передавая по ним только наиболее существенную часть информации, но и для других областей техники, где требуется запомнить большой объем информации, так как при этом экономится значительное число ячеек памяти.

**Первой работой, в которой было введено одно из важнейших понятий теории связи — понятие информации, содержащейся в сообщении,** стала статья [9] американского ученого Р. Хартли, опубликованная в 1928 г. До этой работы само понятие «сообщение» носило расплывчатый характер. В работе Хартли информация, содержащаяся в сообщении, связывалась с последовательным выбором на передающем конце линии связи символов или слов независимо от их смысла. Он показал, что в сообщении из  $N$  символов, выбранных из алфавита, имеющего  $K$  «букв», возможно формирование  $M = K^N$  сообщений. Р. Хартли предполагал, что каждая «буква» выбирается для передачи по каналу связи с одинаковой вероятностью  $1/K$ , и предложил определять информацию, содержащуюся в одном передаваемом символе сообщения, равной  $\log M/N = \log K$ . Идея Р. Хартли использовать логарифмическую меру информации, содержащейся в сообщении, была позже существенно обобщена **К. Шенноном**.

Следует отметить, что характерной чертой современной теории связи является статистический подход к описанию как сообщений, передаваемых по каналу связи, так и действующих в канале связи шумов. Такой подход также начали применять в теории связи до публикации работ К. Шеннона. Одной из первых работ, в которых вопросы помехоустойчивости приема нескольких используемых в те годы на практике методов передачи дискретных сообщений были исследованы с учетом действия в канале связи случайных шумов, явилась докторская диссертация члена-корреспондента АН СССР **В.И. Сифорова** [10]. Многими своими принципами современная статистическая теория связи обязана **Н. Винеру**. Его классическая работа [11] «Ин-

терполяция, экстраполяция и сглаживание временных рядов», выполненная в 1942 г., но рассекреченная и опубликованная только в 1949 г., содержала первую четкую формулировку теории связи как статистической проблемы.

Таковы были основные результаты теории связи, полученные учеными до появления работы К. Шеннона.

## § 2. Создание теории информации

В 1948 г. отдельные разрозненные идеи предшественников были обобщены **К. Шенноном** в его знаменитой работе «Математическая теория связи» [1]. В ней дано четкое математическое описание проблем, стоящих перед теорией связи, и указаны пути их решения. Система связи любого назначения представлена структурной схемой (рис. 4.1), на которой выделены основные функциональные элементы.



Рис. 4.1

Источник на рис. 4.1 представляет собой случайный генератор данных или аналогового стохастического сигнала, а кодер источника осуществляет отображения его выхода в дискретную последовательность символов (обычно двоичную). Назначение кодера источника сообщений — устранение избыточности сообщений, поступающих от источника. Например, если источником сообщений являются буквы алфавита, закодированные равномерным кодом, то, учитывая вероятности их появления в тексте, в кодере источника можно наиболее вероятные буквы перекодировать, как это осуществлено в коде Морзе, короткими кодовыми комбинациями. Простейшим примером кодера аналогового источника служит аналого-цифровой преобразователь, называемый также квантователем; для него декодер источника — цифроаналоговый преобразователь. Мерой достигаемого сжатия данных источника служит скорость, выраженная числом символов (обычно двоичных) в единицу времени, которое необходимо для полного представления и последующего восстановления декодером источника выходной последовательности источника. При этом учитываются статистические характеристики источника и допускается определенный уровень искажений сигнала, поступающего от источника сообщений. Чем выше допустимый уровень искажений, тем экономнее описание сообщения, поступающего от источника. Кодеры и декодеры должны осуществлять взаимно однозначные преобразования таким образом, чтобы получатель сообщений в отсутствие искажений в канале получал тот же сигнал, который действует на входе кодера источника.

Назначением кодера и декодера канала является обеспечение надежной передачи выходной последовательности кодера источника по каналу связи с шумом. На выходе декодера, несмотря на ошибки в приеме сигналов, возникающие из-за действующих в канале связи шумов, должна формироваться с высокой вероятностью та же последовательность, которая поступает на вход кодера канала связи. Этого можно достичь, используя эффективные коды, исправляющие ошибки, которые возникают при передаче сообщений по каналу связи.

По Шеннону, теория связи имеет дело с двумя основными проблемами:

**1. Кодирование источника** таким образом, чтобы среднее число двоичных символов, приходящихся на один символ источника сообщений, было минимальным. Решение этой проблемы позволяет устранить избыточность из сообщения, подлежащего передаче, и достичь высокой эффективности использования канала связи либо объема памяти устройства хранения сообщений.

К. Шенноном, а затем многими другими учеными были установлены теоремы, определяющие скорость создания сообщений источниками с разными статистическими характеристиками. Так, для дискретного источника без памяти, у которого последовательные символы порождаются независимо и могут принимать с вероятностью  $p_i$  значения  $a_i$  ( $i = 1, K$ ), среднее количество информации на один символ на выходе источника задается следующим выражением:

$$H(p_i) = \sum_{i=1}^K p_i \log(1/p_i) \leq \log K. \quad (1)$$

Отметим, что в случае, когда  $p_i = 1/K$ , данная формула совпадает с приведенной выше формулой Р. Хартли. Если  $p_i$  значительно отличаются друг от друга по величине, то  $H(p_i) \ll \log(K)$ , т.е. в данном случае возможно существенное сокращение избыточности сообщений, подлежащих передаче по каналу связи.

К формуле (1) можно прийти с помощью следующих простых соображений, приведенных **К. Шенноном** в своей работе. Они показывают связь подходов К. Шеннона и Р. Хартли к определению количества информации, содержащегося в сообщении. Выбирая достаточно длинную (длиной  $N$ ) последовательность символов, поступающих от источника, можно утверждать, что в любой из таких последовательностей число символов  $a_i$  составит  $Np_i$ . Доля последовательностей, в которых количество символов  $a_i$  отличается от  $Np_i$ , ничтожно мала. Таким образом, число возможных последовательностей длины  $N$ , генерируемых источником, составляет

$$M = N! / \prod_{i=1}^K (Np_i)! \quad (2)$$

Здесь ! — знак факториала. Используя формулу Стирлинга  $k! \approx \sqrt{2\pi k} (k/e)^k$ , из (2) можно получить

$$H(p_i) = \lim_{N \rightarrow \infty} \{\log M / N\}.$$

Функцию  $H(p_i)$  называют энтропией дискретного источника информации без памяти. Одним из методов сокращения избыточности источника сообщений в рассмотренном случае является метод блочного кодирования, когда на выходе кодера источ-