

1 Краткие теоретические сведения

Если сообщения на входе и выходе канала являются дискретными, то такой канал также называют **дискретным**. Обобщённая схема передачи сообщения по дискретному каналу связи (ДКС) показана на рис. 1.



Рис. 1. Передача сообщения по ДКС

Сообщение с выхода дискретного источника (ИС) через кодер (КОД) поступает на вход дискретного канала связи (ДКС). Из-за возникающих в ДКС ошибок сообщение на его выходе может отличаться от сообщения на входе. Соответственно, получателю (ПС) поступает сообщение, которое также может отличаться от исходного. A , B , B' и A' — алфавиты источника, кодера, выхода канала и выхода декодера соответственно с известными совместными вероятностями.

Кодер выполняет функцию согласования источника сообщений с каналом связи. В случаях, когда согласование не требуется, кодер может отсутствовать, и тогда $A = B$ и $A' = B'$. Операция кодирования однозначна, т.е. потерь информации в кодере нет.

ДКС полностью описывается **вероятностями переходов** $P(b'_j|b_i)$ — вероятность принять символ b'_j при передаче символа b_i . Случай $i = j$ соответствует правильной передаче, а все остальные ($i \neq j$) — ошибочной. Если вероятности переходов не зависят от того, какие символы передавались и принимались ранее, такой канал называют **каналом без памяти**.

Симметричным называется ДКС, для которого вероятности ошибочных переходов $P(b'_j|b_i)$ ($i \neq j$) равны между собой, так же как равны между собой и вероятности правильного приёма $P(b'_i|b_i)$ ($i = j$).

m -ичным симметричным каналом без памяти называют канал, объём алфавита на входе и выходе которого равняется m , вероятность ошибочного приёма равняется p , правильного приёма $q = 1 - p$ и вероятность ошибки любого вида одинакова и равна

$$P(b'_j|b_i) = \frac{p}{m-1} \quad (i \neq j).$$

1.1 Информационные характеристики дискретного канала

Среднее количество информации, теряемой при передаче по каналу одного символа, называют **ненадёжностью канала**:

$$H(B|B') = - \sum_{i=1}^m \sum_{j=1}^m P(b_i, b'_j) \log_2 \frac{P(b_i, b'_j)}{P(b'_j)}.$$

Здесь $P(b_i, b'_j)$ — совместная вероятность символов b_i и b'_j .

Ненадёжность канала, как и энтропия дискретного источника, по определению не может быть отрицательной:

$$0 \leq H(B|B') \leq H(B).$$

Равенство $H(B|B') = 0$ достигается при отсутствии в канале ошибок. Равенство $H(B|B') = H(B)$ достигается, когда вход и выход канала статистически независимы, т.е. символы на выходе регистрируются независимо от символов на входе.

Количество информации, переданной по каналу, определяется разностью между количеством информации на входе канала (энтропией сообщения $H(B)$) и количеством информации, потерянной в канале $H(B|B')$:

$$I(B, B') = H(B) - H(B|B') = \sum_{i=1}^m \sum_{j=1}^m P(b_i, b'_j) \log_2 \frac{P(b_i, b'_j)}{P(b_i)P(b'_j)}.$$

Эта формула симметрична относительно b_i и b'_j , поэтому можно записать, что

$$I(B, B') = H(B') - H(B'|B).$$

Здесь $H(B')$ — энтропия на выходе канала, $H(B'|B)$ — **энтропия шума**:

$$H(B'|B) = - \sum_{i=1}^m \sum_{j=1}^m P(b_i, b'_j) \log_2 \frac{P(b_i, b'_j)}{P(b_i)}.$$

Для m -ичного симметричного канала без памяти

$$H(B'|B) = -(1-p) \log(1-p) - p \log \frac{p}{m-1}.$$

Для энтропии шума также соблюдается условие

$$0 \leq H(B'|B) \leq H(B').$$

Если символы входного сообщения распределены равномерно ($P(b_i) = \frac{1}{K}$, K — объём алфавита источника), то энтропия на входе и выходе канала будет одинаковой и максимально возможной:

$$H(B) = H(B') = \log_2 K.$$

В этом случае энтропия шума будет равна ненадёжности канала: $H(B'|B) = H(B|B')$.

Поскольку $H(B|B') \leq H(B)$, то и $I(B, B') \leq H(B)$, т.е. количество информации, переданное по каналу, не может превышать количество информации на его входе. Лишь в предельном случае, когда в канале нет ошибок и выходные последовательности тождественны входным, количество информации на выходе будет равно количеству информации на входе: $I(B, B') = H(B)$.

1.2 Скорость передачи и пропускная способность дискретного канала

Если на вход дискретного канала поступает в среднем V символов в секунду, то можно определить среднюю **скорость передачи** информации по каналу с шумом:

$$I'(B, B') = V \cdot I(B, B') = H'(B) - H'(B|B') = H'(B') - H'(B'|B).$$

Здесь $H'(B) = V \cdot H(B)$ — производительность источника на входе канала, $H'(B|B') = V \cdot H(B|B')$ — ненадёжность канала в единицу времени, $H'(B') = V \cdot H(B')$ — производительность источника, образованного выходом канала, $H'(B'|B) = V \cdot H(B'|B)$ — количество ложной информации, создаваемой помехой в единицу времени.

Предельная возможная скорость передачи информации по заданному каналу с шумом называется **пропускной способностью канала**:

$$C = \max I'(B, B').$$

Для пропускной способности соблюдается условие

$$0 \leq C \leq V \log_2 m,$$

причём $C = 0$, если вход и выход канала независимы, и $C = H'_{\max}(B) = V \log_2 m$, если ошибок нет и энтропия входных символов максимальна.

Пропускная способность m -ичного симметричного канала без памяти

$$C = V \left[\log_2 m + (1 - p) \log_2(1 - p) + p \log_2 \frac{p}{m - 1} \right].$$

Теорема Шеннона

С понятием пропускной способности канала непосредственно связана одна из фундаментальных теорем теории информации, известная как теорема Шеннона об оптимальном кодировании.

Теорема Шеннона: Если производительность источника сообщений $H'(A)$ меньше пропускной способности обслуживающего его канала, т.е. $H'(A) < C$, то существует способ оптимального кодирования и декодирования, при котором вероятность ошибки может быть сколь угодно малой:

$$P(\text{ош}) < 2^{-T[C-H'(A)]}.$$

Если же $H'(A) > C$, то таких способов не существует.

Суть теоремы заключается в следующем. При повышении скорости передачи информации возрастают потери в дискретном канале связи $B \rightarrow B'$. Однако, применяя оптимальное кодирование, можно свести потери в канале $A \rightarrow A'$ к сколь угодно малой величине.

2 Домашнее задание

Определить ненадёжность, энтропию шума, количество переданной информации и пропускную способность двоичного симметричного канала при вероятности ошибки $p = abc \cdot 10^{-4}$ (abc — три последние цифры номера студенческого билета). При этом считать, что кодирование не применяется ($m = K$), выходные символы источника сообщений равновероятны, независимы и поступают на вход канала со скоростью $V = abc \cdot 10^3$ симв/с.

3 Указания к выполнению работы

Перед началом работы ознакомьтесь с интерфейсом пользователя: нажмите **F1** или выберите в главном меню «Помощь → Руководство пользователя».

3.1 Цель работы

Целью работы является анализ информационных характеристик дискретных каналов.

3.2 Общие замечания

Для получения достоверных результатов требуется большой объём статистики, поэтому при снятии зависимостей не рекомендуется брать

короткие тексты (длиной менее 100 тысяч символов, для медленных машин — 20–30 тысяч символов).

В работе рассматривается m -ичный симметричный канал без памяти с объёмом алфавита, равным объёму алфавита источника: $m = K$. Кодирование не применяется и, следовательно, $H(B) = H(A)$, $H(B') = H(A')$, $H(B|B') = H(A|A')$, $H(B'|B) = H(A'|A)$ и $I(B, B') = I(A, A')$. Объём алфавита текста на русском языке $K = 32$.

3.3 Исследование информационных характеристик дискретного канала с известными параметрами

Снимите зависимость информационных характеристик m -ичного симметричного канала без памяти от вероятности ошибки p .

1. Выберите канал с известными параметрами (известный канал).
2. Сгенерируйте случайный текст на русском языке («Файл → Случайный текст» или Ctrl+R) с распределением вероятностей, соответствующим посимвольной модели русского языка. Длину текста установите в диапазоне от 100000 до 1000000 символов (20–30 тысяч для медленных машин).
3. Заполните табл. 1.

Таблица 1

p	$H(A)$	$H(A')$	$H(A A')$	$H(A' A)$	$I(A, A')$
	[бит/симв]				
0					
0,1					
0,2					
0,3					
0,4					
0,5					
0,6					
0,7					
0,8					
0,9					
1					

4. Повторите то же самое для случайного текста на русском языке с равномерным распределением вероятностей и для случайного двоичного текста с равномерным распределением вероятностей.

По полученным таблицам постройте графики зависимостей $H(A|A')$, $H(A'|A)$ и $I(A, A')$ от p . Сделайте выводы.

3.4 Исследование пропускной способности дискретного канала со случайными параметрами

Определите опытным путём пропускную способность m -ичного симметричного канала без памяти с заранее неизвестными (случайно выбранными) параметрами. **Внимание!** Параметры канала выбираются случайно при каждом запуске программы, поэтому не перезапускайте её в процессе выполнения данного пункта.

1. Выберите канал со случайными параметрами (случайный канал).
2. Сгенерируйте случайный текст на русском языке («Файл → Случайный текст» или Ctrl+R) с распределением вероятностей, соответствующим посимвольной модели русского языка. Длину текста установите в диапазоне от 100000 до 1000000 символов (20–30 тысяч для медленных машин).
3. Пропускная способность — это максимальная скорость передачи информации по данному каналу: $C = \max I'(A, A')$. Для отыскания максимума $I'(A, A')$ сначала определите его примерное положение, увеличивая V грубыми шагами (500–1000 [симв/с]), затем уточните его, изменяя V мелкими шагами (50–100 [симв/с]) в области максимума. Все получаемые в процессе эксперимента значения заносите в таблицу (табл. 2).

Таблица 2

V [симв/с]	p^*	$H'(A)$	$H'(A A')$	$I'(A, A')$
		[бит/с]		
100				
...				

4. Укажите в таблице полученное значение пропускной способности: $\max I'(A, A') = C$.
5. Повторите то же самое для случайного текста на русском языке с равномерным распределением вероятностей и для случайного двоичного текста с равномерным распределением вероятностей.

Сравните полученные результаты. Сделайте выводы.

4 Содержание отчёта

1. Название и цель работы.
2. Выполненное домашнее задание.
3. Таблицы, графики и выводы по всем пунктам работы.
4. Общий вывод по результатам лабораторной работы.

Примечание: общий вывод не должен быть перефразировкой целей работы, а должен содержать обобщение (но не копию!) всех выводов, сделанных по каждому пункту в отдельности.

5 Контрольные вопросы

1. Какой канал называется дискретным? Как выглядит схема передачи сообщения по дискретному каналу?
2. Какими параметрами описывается дискретный канал связи? Какой канал называется каналом без памяти? симметричным каналом?
3. Дайте определение ненадёжности канала. В каких пределах изменяется ненадёжность?
4. Как определяется количество информации, переданной по каналу? В каких пределах оно может изменяться?
5. Дайте определение энтропии шума. Чему равна энтропия шума в m -ичном симметричном канале без памяти?
6. При каком условии ненадёжность канала будет равняться энтропии шума?
7. Как определяется скорость передачи информации по дискретному каналу?
8. Дайте определение пропускной способности канала. Чему равна пропускная способность m -ичного симметричного канала без памяти?
9. О чём говорит теорема Шеннона?
10. Чему равняется энтропия шума в двоичном симметричном канале без памяти с вероятностью ошибки $p = 0,5$?

11. Чему равняется количество информации, переданной по двоичному симметричному каналу без памяти с вероятностью ошибки $p = 1$?
12. При каком значении вероятности ошибки количество информации, переданной по m -ичному симметричному каналу без памяти, равняется нулю?
13. Как зависит пропускная способность канала от скорости передачи информации?

Литература

1. Кловский Д. Д. Теория электрической связи. — М.: Радиотехника, 2009. — 648 с.
2. Теория электрической связи: учебник для вузов / А. Г. Зюко, Д. Д. Кловский, В. И. Коржик, М. В. Назаров; Под ред. Д. Д. Кловского. — М.: Радио и связь, 1998. — 432 с.
3. Шеннон К. Работы по теории информации и кибернетике. — М.: Издательство иностранной литературы, 1963.
4. Яглом И. М., Яглом А. М. Вероятность и информация. — М.: Наука, 1973.