ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТИ В ВЫЧИСЛЕНИИ ИНФОРМАЦИОННЫХ ПОТЕРЬ ПРИ ПЕРЕДАЧИ СООБЩЕНИЙ ПО КАНАЛАМ СВЯЗИ С ШУМАМИ

Ганиева Зулфия Самиевна - Самаркандский институт экономики и сервиса, ассистент

Аннотация. В статье рассмотрено применение некоторых теорем теории вероятностей для процессов, присущих передаче сообщений по каналам связи частности приведены задачи на энтропию.

Ключевые слова: теория информатизации, теория вероятностей, условная энтропия, условная и безусловная вероятность, матрица, среднее количество информации, вероятность связи.

Annotation. The application of some theorems of probability theory for processes inherent in the transmission of messages over communication channels is considered; in particular, entropy problems are presented.

Key words: information theory, probability theory, conditional entropy, conditional and unconditional probability, matrix, average amount of information, communication probability.

Одной из задач теории информации является отыскание наиболее экономных методов кодирования, позволяющих передать заданную информацию с помощью минимального количества символов. Эта задача решается как при отсутствии, так и при наличии искажений (помех) в канале связи [1].

Потери информации в каналах связи с шумами описывают с помощью условной энтропии и энтропии объединения. Если помех нет или их уровень низкий настолько, что они не могут уничтожить сигнал или имитировать сигнал в отсутствие передачи, то при передаче a_i будет получен сигнал b_j, соответствующий переданному сигналу. События A и B жестко связаны, при этом условная вероятность максимальна $P(b_j/a_i) = 1$, и условная энтропия $H(A/B) = 0$, так как log $P(b_j/a_i) = 0$. Для этого случая количество информации в принятом ансамбле сообщений B, равно энтропии передаваемых сообщений ансамбля A

$$I(B,A) = H(A)$$

Если уровень помех высок, то любой из принятых сигналов b_j может соответствовать любому переданному сигналу a_i, статистическая связь между переданными и принятыми сигналами отсутствует. Поэтому вероятности $P(a_i)$ и $P(b_j)$ являются вероятностями независимых событий и

$$P(b_j/a_i) = P(b_j); \qquad P(a_i/b_j) = P(a_i).$$

$$H(A/B) = -\sum_i \sum_j P(b_j) P(a_i/b_j) log P(a_i/b_j) = ¿$$

$$\text{¿}-\sum_i \sum_j P(b_j)P(a_i) logP(a_j) = \sum_j P(b_j)H(A) = H(A)$$

Так как $\sum_j P(b_j)=1$, условная энтропия равна безусловной, а количество информации, содержащееся в В, относительно А равно нулю.

$$I(A,B) = H(A) - H(A/B) = 0$$

Информационные характеристики реальных каналов связи лежат между этими двумя предельными случаями. При этом потери информации при передаче k символов по данному каналу связи

$$\Delta I = kH(A/B)$$

Из-за помех часть информации искажается, однако между переданными и принятыми сообщениями существует статистическая взаимосвязь. Это позволяет описывать информационные характеристики реальных каналов связи с помощью энтропии объединения статистически зависимых событий.

Так как

$$H(A,B) = H(A) + H(B/A) = H(B) + H(A/B),$$

то потери в канале связи могут быть учтены с помощью энтропии объединения следующим образом

$$I(B,A) = H(A) + H(B) - H(B,A).$$

Если использовать условную энтропию, то получим

$$I(B,A) = H(A) - H(A/B) = H(B) - H(B/A).$$

Для вычисления среднего количества информации, содержащегося в принятом ансамбле сообщений В относительно переданного ансамбля А в условиях действия помех, пользуются следующими выражениями

$$I(B,A) = \sum_i \sum_j P(a_i) P(b_j/a_i) \log \frac{P(b_j/a_i)}{P(b_j)} \qquad (**)$$

$$I(A,B) = \sum_i \sum_j P(b_j) P(a_i/b_j) \log \frac{P(a_i/b_j)}{P(a_i)} \qquad (**)$$

Для вычислений часто применяют выражения:

$$I(A,B) = \sum_j P(b_j) \sum_i [P(a_i/b_j) logP(a_i/b_j) - P(a_i/b_j) logP(a_i)],$$

$$I(B,A) = \sum_j P(a_i) \sum_i [P(\text{¿}a_i) logP(b_j/a_i) - P(b_j/a_i) logP(b_j)]$$

$$I(A,B) = I(B,A) = \sum_i \sum_j P(a_i,b_j) logP(a_i,b_j) - \sum_i \sum_j P(a_i,b_j) logP(a_i) P(b_j)$$

Для полного описания канала связи необходимо задать: канальную матрицу вида P(ai/bj) и безусловные вероятности вида P(bj) или канальную матрицу вида P(bj/ ai) и безусловные вероятности P(ai), или канальную матрицу вида P(ai, bj). В последнем случае сумма значений матрицы по столбцам дает безусловные вероятности вида $P(b_j)(\sum_j P(b_j)=1)$, а сумма по строкам дает безусловные вероятности вида $P(a_i)(\sum_i P(a_i)=1)$. Условные вероятности могут быть найдены из выражений

$$P(a_i/b_j)=\frac{P(b_j,a_i)}{P(b_j)} \qquad (*)$$

$$P(b_j/a_I)=\frac{P(a_I,b_J)}{P(a_i)} \qquad (*)$$

Это (*) частные условные энтропии, индекс i – выбран для характеристики произвольного состояния источника сообщений А, j – выбран для характеристики произвольного состояния адресата В. (**) общая условная энтропия сообщения В относительно сообщения А. характеризует количество информации, содержащееся в любом символе алфавита, и, определяется усреднением по всем символам, то есть по всем состояниям с учетов вероятности появления каждого из состояний, и равна сумме вероятностей появления символов алфавита на неопределенность которая остается после того, как адресат принял сигнал.

Зная условные и безусловные вероятности, можно найти энтропии *H(A), H(B), H(A/B), H(B/A).* Если уровень помех настолько высок, что с равной вероятностью можно ожидать переход любого символа источника сообщения в произвольный символ первичного алфавита, то энтропия канала связи будет равна log m, а количество информации I = H(A) – log m≤ 0, значение I может отрицательной величиной, что означает, что канал связи вносит дезинформацию.

Пример. Канал связи задан следующей канальной матрицей

$$P(b,a)=\begin{vmatrix} 0{,}98 & 0{,}01 & 0{,}01 \\ 0{,}1 & 0{,}75 & 0{,}15 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{vmatrix}$$

Вычислить среднее количество информации, которое переносится одним символом сообщения, если вероятности появления символов источника сообщений равны $P(a_1) = 0{,}7$; $P(a_2) = 0{,}2$; $P(a_3) = 0{,}1$. Определить информационные потери при передаче сообщения из 400 символов алфавита a_1, a_2, a_3. Вычислить количество принятой информации.

Решение. Энтропия источника сообщений

$$H(A)=\sum_{i=1}^{m} P_i \log P_i=-(0{,}7\log 0{,}7+0{,}2\log 0{,}2+0{,}1\log 0{,}1)=1{,}1568 \text{ бит/символ}$$

Общая условная энтропия

$$H(A)=-\sum_i P_i \log(a_i)\sum_i P\left(\frac{b_{JI}}{a_i}\right)\log\left(\frac{b_{JI}}{a_i}\right)=-[0{,}7*(0{,}98\log 0{,}98+2*0{,}01\log 0{,}01)+0{,}2*(0{,}75\log 0{,}75+$$

бит/символ.

Потери в канале связи

ΔI = kH(B/ A) = 400*0,473 = 189,5 бит.

Энтропия приемника

$$H(B)=\sum_{j=1}^{m} P(b\text{¿¿}j)\log P(b\text{¿¿}j)\text{¿¿}$$

$H(b_1)=\sum_{i} P(a\text{¿¿}i)P(b_1/a_i)=\text{¿}P(a\text{¿¿}1)P(b_1\backslash a_1)+P(a\text{¿¿}2)P(b_1\backslash a_2)+P(a\text{¿¿}3)P(b_1\backslash a_3)=0{,}726\text{¿¿¿¿}$

$H(b_2)=\sum_{i} P(a\text{¿¿}i)P(b_2\backslash a_i)=\text{¿}P(a\text{¿¿}1)P(b_2\backslash a_1)+P(a\text{¿¿}2)P(b_2\backslash a_2)+P(a\text{¿¿}3)P(b_2\backslash a_3)=0{,}187\text{¿¿¿}$

$H(b_3)=\sum_{i} P(a\text{¿¿}i)P(b_3\backslash a_i)=\text{¿}P(a\text{¿¿}1)P(b_3\backslash a_1)+P(a\text{¿¿}2)P(b_3\backslash a_2)+P(a\text{¿¿}3)P(b_3\backslash a_3)=0{,}087\text{¿¿¿}$

$H(b_1)+H(b_2)+H(b_3)=1$

$H(B)=\sum_{j=1}^{3} P(b\text{¿¿}j)\log P(b\text{¿¿}j)=1{,}094\text{¿¿}$ *бит/символ*

Среднее количество принятой информации

$I = k\,[H(B) – H(B/A)] = k\,H(B) - \Delta I = 400 * 1{,}094 – 189{,}5 = 248{,}1$ *бит.*

Таким образом, понятия зависимых событий, теоремы об условной вероятности и полной вероятности применяются (в соответствующих обозначениях) в передаче сообщений на расстоянии, что делает весьма важным изучении теории вероятностей в техническом ВУЗе.

Литература:

1. Вентцель Е.С. Теория вероятностей: Учебник для вузов.— 6-е изд. стер.— М.: Высш. шк., 1999.— 576 с.
2. Теория информации: методические указания / Е.В. Бурькова; Оренбургский гос. ун-т. – Оренбург: ОГУ, 2018. – 50 с.
3. Цымбал В.П. Задачник по теории информатизации и кодированию.- изв. Вища школа.-Киев, 1976.-276 с.
4. Зверева Е.Н., Лебедько Е.Г. Сборник примеров и задач по основам теории информации и кодирования сообщений. Методические указания. – СПб: НИУ ИТМО, 2014. – 76 с.