

УДК 621.391

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ В
ИНФОРМАЦИОННЫХ СИСТЕМАХ

Дулесов А.С.

Хакасский государственный университет им. Н.Ф. Катанова, Абакан

Подробная информация об авторах размещена на сайте

«Учёные России» - <http://www.famous-scientists.ru>

Рассматриваются вопросы соотношения энтропии и информации. Даны математические выражения для определения объема информации закрытых систем.

Эволюция ноосферных процессов непосредственно связана с наличием информационной среды в которой находится огромное количество объектов различной природы с наличием в них информационных потоков. Поэтому открытыми для исследований остаются вопросы рациональ-

ного использования информации, построения информационных систем и др.

Чтобы ответить хотя бы на малую часть этих вопросов затронем понятие «количество информации». Оно дано Шенноном [1] и выглядит следующим образом:

$$I = - \sum p_i \ln p_i. \tag{1}$$

В (1) суммирование производится по всем символам, а p_i означает вероятность появления символа с номером i . Здесь используется натуральный логарифм, поэтому единица измерения информации получила название «нат».

Используя (1) не только для одного знака (например, буквы алфавита), для текста из N знаков эту информацию следует увеличить в N раз.

Если перейти к измерению информации в битах, используется выражение

$$I_6 = - \log_2 p_N, \tag{2}$$

в котором p_N – средняя вероятность появления в потоке информации различных текстов, то переход от натов к битам возможен через выражение:

$$I_6 = I / \ln 2 \approx 1.44 I. \tag{3}$$

Следовательно, информация может быть измерена в натах и битах при условии соблюдения перехода (3).

Информация (1) играет большую роль в информационных процессах и системах, в которых потоки информации определяют упорядоченность связей, их надежность функционирования и эффективность управления. Подчеркнем, что речь идет именно о числовом выражении информации.

Говоря об информации, следует сделать упор на понятие энтропии, как фундаментального понятия физики. С целью установления связи между информацией и энтропией обратимся к работам

Хазена А. М. [2] и Кадомцева Б.Б. [3], что позволит провести параллели между процессами, протекающими в технических и информационных системах.

Не вдаваясь в математические формулировки описывающие энтропию и в наличие количества аксиом термодинамики, можно сказать: энтропия – есть функция состояния системы. Эта формулировка наиболее логично соотносится со вторым началом термодинамики. В дополнение к данной формулировке следует добавить еще две: 1) изменения энтропии в неадиабатических процессах дискретны; 2) в адиабатических процессах энтропия или

увеличивается или остается неизменной (дана Энштейном).

Согласно [3] связь энтропии и информации записывается как

$$S + I = const, \quad (4)$$

где S – энтропия.

Выражение (4) показывает: если одна из составляющих будет последовательно уменьшаться (например, энтропия), то информация будет увеличиваться. Справедливо и обратное утверждение.

Чем больше известно о системе, тем меньше её энтропия. То же можно сказать и об информации: устранение неопределенности в системе (например, за счет рационального управления) приводит к снижению энтропии-информации. Наличие хаоса в системе свидетельствует об отсутствии порядка и наличии максимума информации.

Следует отметить, что появление информации во внешнем мире невозможно без возрастания энтропии внешнего окружения на величину, не меньшую DI . Следовательно, информация имеет информационную емкость, может быть усвоена (например, приборами, средствами вычислительной техники) за счет появления во внешнем окружении дополнительных источников информации (наличие теплового движения) с возрастанием энтропии не меньшим, чем усвоено информации.

Согласно (4) сумма S и I постоянна. Если рассматривать некоторое информационное пространство в котором находится элементарная частица информации (например, буква в книге), то при этом энтропия $S = 0$, а информация достигнет максимального значения $I_{max} = -\ln p_{min} = \ln(1/p_{min})$, поскольку вероятность найти частицу информации во всем информационном пространстве очень мала. Для того, чтобы найти и зафиксировать (распознать) частицу, то есть усвоить информацию (например, с помощью приборов или средств вычислительной техники) потребуется произвести не меньшее количество энтропии в приборе или за его пределами. Если включить механизм усвоения информации, энтропия начнет возрастать, а информация будет стремиться к нулю $I_{min} = \lim_{p \rightarrow 1} (\ln(1/p)) \rightarrow 0$, так как вероят-

ность распознать частицу будет увеличиваться.

Состояние с максимальной энтропией – наиболее устойчивое состояние для изолированной системы (Э. Ферми). При этом состоянии энтропии информация будет минимальной, свидетельствующей о полной определенности информации в системе. Эти утверждения являются следствием процедуры нормировки энтропии и информации, то есть установления связи количества энергии с количеством элементов системы и с величиной её энтропии-информации.

Соотношение (4) свидетельствует о том, что между тепловым и информационным процессами существует взаимосвязь. Следовательно, вопросы аксиоматики термодинамики распространяются на любой вид информации, представленной в больцмановском виде, $S = \ln G$. Природа стремится к переходу от менее вероятных состояний к более вероятным (Л. Больцман). Тем самым энтропия физической системы связана с распределением вероятностей. Здесь G – число возможных микросостояний элементов, которые отвечают одному и тому же макросостоянию системы.

Величины S и I равны лишь формально, поскольку I отвечает максимальной информации только одного состояния, а S определена по множеству всех состояний. Энтропия и информация замкнутой системы оказываются взаимными по отношению друг к другу: «забывание» информации приводит к возрастанию энтропии.

Если рассматривать замкнутую систему, то согласно второго начала термодинамики энтропия не может убывать со временем (аксиома Эйнштейна). Следовательно, информация будет менять знак на противоположный, что свидетельствует о её забывании. Новая часть информации, т.е. её осмысленная фиксация может происходить только за счет дополнительного стирания куска информации, если данный

процесс происходит без внесения добавочной информации извне. В замкнутой системе порядок образуется только за счет уничтожения другого порядка.

Тем самым рассматривая закрытую информационную систему необходимо соблюдать условие баланса между информацией, идущей от источников и информацией, принимаемой приемниками.

Рассмотрев взаимосвязь между S и I отметим, что аксиома: энтропия – есть функция состояния системы, распространяет аксиомы термодинамики на любой вид информации, представленной в больцмановском виде, даже если она абстрактная.

Когда речь идет о распределении потоков информации в системе, то рассмотренные выше аксиоматические связи между S и I позволяют свидетельствовать о присутствии меры количества информации, её ёмкости (информационная емкость – максимальное количество информации, которое может содержаться или быть обработано в данном объекте или канале), суммировании, нуле отсчета (аксиоматическое свойство, изложенное В. Нернстом), а также смене знака, если принимается во внимание направление её движения.

Опираясь на [2] подчеркнем ниже следующее:

I. Энтропия-информация – есть иерархическая функция состояния системы, которая определена в фазовом пространстве для заданных признаков и условий элементов системы. Вычисляется через выражение (1), которое является мерой количества информации (мерой фазового пространства) в пределах заданных признаков и условий для наиболее вероятного состояния системы из многих элементов.

II. Энтропия-информация достигает максимума вероятности состояния системы, которая нормирована по отношению к энергии и к числу элементов системы.

III. Энтропию-информацию порождает процесс синтеза информации, то есть запоминание случайного выбора. В про-

цессе выбора критерии запоминания (устойчивости) зависят от экстремумов энтропии-информации и её производства. Случай синтеза информации в системе описывает принцип максимума производства энтропии-информации, определяет условия разрушения равновесия и перехода к следующей ступени иерархии роста энтропии-информации.

IV. Энтропия-информация подлежит суммированию при входящих в её определение признаках с учетом уравнений связи их между собой. Для любого признака существует свой нуль отсчета, который зависит от содержания признака. Энтропия-информация – положительно определённая переменная. Однако существование нулей отсчета для признаков разрешает в конкретных задачах использовать информацию с отрицательным знаком.

V. Энергия – функция состояния системы. Энергию можно представить как сумму разных её форм. Существует множество форм энергии, например, в более общем виде – информационная энергия или тепловая энергия, которая выражается произведением температуры на энтропию. Сумма энергий одной формы от разных источников в системе равна нулю, то есть общее количество энергии, подведенное к материальному объекту, равно количеству энергии отведенного от данного объекта. Тепловая энергия связана с температурой, которая есть обратный масштаб измерения времени в замкнутой системе. Время в замкнутой системе обратимо. Сохранение величины суммы форм энергии (закон сохранения энергии) относится к замкнутым системам. Энергия системы изменяется в результате взаимодействия системы с окружающей средой. Замкнутая система в любой точке своей границы находится в статическом и динамическом равновесии с окружением.

Количество информации содержащейся в одном объекте равна сумме количеств информации, которые были получены им от n независимых i -х источников с N_i числом возможных сообщений, то есть

$$I = \sum_{i=1}^n \log N_i \quad (5)$$

Выражение (5) (формула Хартли) можно применить, если рассматривать потоковый процесс, когда частицы информации пробегают по каналу связи с равной вероятностью и статистически не зависимы, то есть присутствует равномерное

распределение частиц на временном интервале.

В случае, когда источник i передает информацию в виде частиц (сигналов) k различных типов, следует обратиться к формуле Шеннона:

$$I = -N \sum_{j=1}^k p_j \log p_j, \quad (6)$$

где p_j – вероятность появления сигнала j -го типа; N – общее число сигналов от источника.

Приемник может получать сигналы от многих источников и различных типов. В данном случае приемник будет принимать информацию в объеме:

$$I = -\sum_{i=1}^n \sum_{j=1}^k \log N_{ij}$$

$$I = -\sum_{i=1}^n N_i \sum_{j=1}^k p_{ij} \log p_{ij}$$

или (7)

где N_{ij} – число однотипных сигналов идущих от i -го источника j -го типа; p_{ij} – вероятность появления сигнала от i -го источника j -го типа.

чаемых приемниками, возможное направление передачи информации в системе.

Наличие в информационной системе источников и приемников, их емкостей, а так же каналов с возможными направлениями перемещения информации позволяет составить сетевую модель. Моделируя на графе потоковые процессы можно определять загруженность каналов, количество документов различного рода полу-

СПИСОК ЛИТЕРАТУРЫ:

1. Шеннон К. Работы по теории информации и кибернетики. – М.: ИЛ. 1963.
2. Хазен А. М. Разум природы и разум человека. – М.: НТЦ Университетский. 2000. – 232 с.
3. Кадомцев Б. Б. Динамика и информация // УФН. 1994. Т. 164. №5. С. 449-530.

QUANTITATIVE APPRECIATION OF INFORMATION IN INFORMATION SYSTEMS

Dulesov A.S.

Katanov State University of Khakassia, Abakan

Entropies' and information correlation questions are considered in this paper. Mathematical expressions for information volume determination of internal systems are given here.