

УДК 621.390

ОПРЕДЕЛЕНИЕ ДЛЯ ПРОСТЕЙШЕЙ СТРУКТУРЫ ТЕХНИЧЕСКОЙ СИСТЕМЫ КОЛИЧЕСТВА ИНФОРМАЦИОННОЙ ЭНТРОПИИ ПОСРЕДСТВОМ ЕЁ НОРМИРОВКИ

Дулесов А.С., Кондрат Н.Н.

ФГОУ ВПО «Хакасский государственный университет им. Н.Ф. Катанова», Абакан,
e-mail: ninok@khsu.ru

Рассматривается структура технической системы в простейшем виде. В ней каждый элемент может находиться в одном из двух противоположных состояниях, к каждому из которых приписывается вероятность. Даются пояснения и классические математические выражения, с применением которых можно определить общее количество информации, содержащейся в структуре системы. С тем, чтобы разделить информацию, относящуюся к каждому из противоположных событий, составлена упрощенная математическая модель. Она включает в себя свойства, присущие энтропийным характеристикам, и математические выражения для определения энтропии качественного состояния системы. Важным условием выработки математических выражений для определения количества информации явилась операция её нормировки. Применение нормировки позволяет находить количество энтропии непосредственно по каждому из противоположных событий. Тем самым удалось для двух взаимосвязанных элементов разделить информацию о структурном содержании системы по качественному признаку.

Ключевые слова: мера информации, информационная энтропия, нормировка количества информации, структура технической системы

DETERMINATION FOR THE SIMPLEST STRUCTURE OF TECHNICAL SYSTEM OF QUANTITY OF INFORMATION ENTROPY BY MEANS OF ITS NORMALIZATION

Dulesov A.S., Kondrat N.N.

Katanov Khakass State University, Abakan, e-mail: ninok@khsu.ru

The structure of technical system in the elementary look is considered. There each element can be in one of two opposite statuses, to each of which the probability is attributed. Explanations and classical mathematical expressions using which it is possible to define total quantity of the information containing in structure of system are given. To separate the information relating to each of opposite events the simplified mathematical model is made. It includes properties inherent in entropy characteristics and mathematical expressions for determination of an entropy of a qualitative status of system. An important condition of framing of mathematical expressions for determination of amount of information was operation of its normalization. The amount of entropy directly on each of opposite events allows to find application of normalization. Thereby it was succeeded to separate information on the structural maintenance of system on a qualitative character for two interdependent elements.

Keywords: information measure, information entropy, normalization of amount of information, structure of technical system

При использовании количества информации в процессе оценки состояния технического объекта желательно определить с понятием «информация». Существует множество определений, относящихся к данному понятию. Например, в [9] информация определяется, как результат выбора – «запомненный выбор». Если обратиться к классической теории информации, в основе которой лежат работы К. Шеннона, то в ней информация определяется как мера или количество неопределённости (информационной энтропии). Само же понятие информации К. Шеннон связывает со снятием неопределённости или с результатом выбора из возможных альтернатив. Как отмечено в [8], информация тесно связана со структурой объектов, поэтому любое их взаимодействие сопровождается изменением зафиксированной в ней информации. Например, структура технической системы

в результате взаимодействия между её элементами меняет свое содержание, что свидетельствует об изменении информации, количество которой можно определить через её меру. Следовательно «результат выбора» (через количественную меру информации) уместен в задачах выбора наиболее информативной структуры из альтернатив.

Отмечая упомянутые в работах [8, 9] и К. Шеннона определения информации, их значимость будет обоснована далее при определении количества информации, содержащейся в простейшей структуре технической системы.

Структура системы и состояние её элементов. Оценка состояния технической системы (например, трубопроводных систем, систем электроснабжения, компьютерных сетей и др.) обусловлена задачами (например, проектирования, эксплуатации), в которых вероятностное состояние струк-

туры можно оценить посредством применения теорий вероятности, надежности и информации.

Учет состояний элементов системы и связей между ними, расчет количества присущей ей информации, позволяет на основе конечных результатов осуществить выбор из имеющихся альтернатив. При этом информацию целесообразно рассматривать как процесс, взаимосвязь, отношение между элементами системы. При этом важна та информация, которая указывает на наличие отклонений параметров системы от заданных условий (или значений) под воздействием сторонних факторов [1].

На практике важно уметь численно оценивать степень неопределенности разнообразных состояний (структурных изменений) системы чтобы иметь возможность сравнить их между собой. При этом энтропия является важным показателем неопределенности, определение величины которой может служить мерой структурного состояния системы. Рост энтропии во времени означает развитие хаотических процессов в системе, что свидетельствует о старении элементов системы, понижении уровня надежности, отсутствии должного управления и организационного поведения [5].

Далее в качестве примера определения меры информации рассмотрим структурную надежность системы. В ней каждому элементу присущи два противоположных дискретных состояния: работоспособное и неработоспособное. Если нас не интересует время пребывания элемента в том или ином состоянии, то общее число всевозможных состояний элементов будет равно $N = k^n$, где $k = 2$ – число возможных дискретных состояний элемента, n – количество элементов в рассматриваемой системе. Данное выражение лежит в основе выражения Хартли для определения меры количества информации в битах: $I = \log_2 N$.

Условие о наличии противоположных и равновероятных состояний позволяет переписать это выражение в иной форме:

$$I = \log_k N = \log_k N_1 + \log_k N_2 = I_1 + I_2, \quad (1)$$

где N_1 и N_2 – количество состояний относящихся к работе и отказу элементов системы соответственно.

Из-за наличия противоположности состояний (работа и отказ), в (1) информация разделена на качественные составляющие. Тем самым учет качественного признака направлен на разделение информации по составляющим с целью последующего выбора. Поэтому процесс определения количества информации должен идти отдельно по каждому из противоположных состо-

яний, поскольку при их слиянии утрачивается смысл оценки уровня надежности через меру информации [2]. Поскольку важнейшей практической задачей является сохранение высокого уровня надежности элемента или системы, то на практике вероятность работоспособного состояния всегда будет выше противоположной вероятности. По этой причине выражение (1) при учете в нем равновероятных состояний малопригодно и может быть использовано только для грубых оценочных расчетов. Устранение данного недостатка приводит нас к необходимости использования формулы Шеннона (при наличии многообразия соотношения противоположных состояний элемента) в следующем виде:

$$I = -(p \log_2 p + q \log_2 q), \quad (2)$$

где, p и q – вероятность нахождения элемента в работоспособном и неработоспособном состоянии, $p + q = 1$. Из (2) вытекает формула Шеннона для n элементов:

$$I = -[\sum_{i=1}^n (p_i \log_2 p_i) + \sum_{i=1}^n (q_i \log_2 q_i)]. \quad (3)$$

Однако выражение (3) малопригодно, поскольку составляющие информации не отражают наличия взаимосвязей между элементами, то есть совместных событий (состояний), присущих элементам. По этой причине следует построить математическую модель для того, чтобы вычислить количество информации о состоянии самой системы с учетом вероятностного появления в ней совместных и условных событий.

Далее предложим к рассмотрению модель определения информационной энтропии (как меры неопределенности) с учетом качественного признака, обусловленного наличием противоположных состояний элементов.

Упрощенная модель определения энтропии качественного состояния простейшей системы. При построении модели под простейшей системой будем понимать такую, структура которой состоит из двух элементов соединенных последовательно, либо параллельно. Даже в такой системе, на первый взгляд кажущейся простейшей, при сопоставлении последовательных и параллельных структур, количество вероятных состояний резко отличается между собой. Например, в работах [6, 7] отмечается существенное отличие в определении энтропии для последовательных и параллельных структур из-за многообразия комбинаций их состояний. Если число элементов увеличить, то резко возрастет и количество состо-

яний, что делает работу по определению энтропии практически невозможной.

При построении математических функций энтропии модели выделим ряд общих свойств:

1) неопределенность состояний структуры системы выражается через вероятностные характеристики системы, полученные на основе статистики;

2) энтропийная характеристика системы зависит от размерности пространства состояний системы и числа элементов системы;

3) энтропийная характеристика разделяет вероятностные состояния системы на два противоположных состояния.

В основу для построения функций энтропии положим результаты исследований, представленные в работах [3, 4]. В них даны математические выражения, с помощью которых величина энтропии наложения и пересечения двух взаимосвязанных событий определяется через геометрическую площадь, образованную, с одной стороны, вероятностью события, с другой – частной информацией, присущей событию.

Предварительно рассмотрим энтропийные характеристики $H(p)$ двух независимых событий с вероятностями p_1 и p_2 :

– совместная энтропия (с учетом совместной вероятности $p_1 p_2$),

$$H(p_1 p_2) = -p_1 p_2 \log_2 p_1 - p_1 p_2 \log_2 p_2 = p_2 H(p_1) + p_1 H(p_2), \quad (4)$$

где $H(p_1) = -p_1 \log_2 p_1$, $H(p_2) = -p_2 \log_2 p_2$;

– условная энтропия (с учетом условных вероятностей p_1/p_2 и p_2/p_1),

$$H(p_1/p_2) = H(p_1) - p_1 H(p_2); \quad (5)$$

$$H(p_2/p_1) = H(p_2) - p_2 H(p_1). \quad (6)$$

$$H(p_1/p_2) + H(p_2/p_1) = H(p_1) + H(p_2) - H(p_1 p_2) = p_1 \log_2 p_1 + p_2 \log_2 p_2 - p_2 H(p_1) - p_1 H(p_2). \quad (7)$$

Выражения (4)–(7) справедливы при условии $p_1 \neq p_2$. Однако они не предусматривают разделение энтропии по качественному признаку. Чтобы устранить этот недостаток (и тем самым выполнить третье свойство искомого выражений), введем условие нормировки меры неопределенности.

Нормировка показателей при построении модели. Прежде дадим пояснения в необходимости нормировки. Рассматривая элемент с двумя состояниями по выражению (1) и с двумя равновероятными состояниями $p = q = 0,5$ по (2), их элементарное сравнение событий (означающее переход от одного состояния к другому)

позволяет получить 1 бит максимума информации. Все остальные элементарные сочетания противоположных событий с вероятностями $p + q = 1$ дают суммарную информацию меньшую единицы. Придерживаясь условия $p_i + q_i = 1$ и $p_i = q_i = 0,5$, согласно (3) максимальное количество информации, присущее структуре системы будет равно n , то есть числу её элементов. Тем самым выделим ещё одно свойство информации: *максимум информации элементарных сравнений противоположных равновероятных событий по каждому из элементов системы равен числу её элементов.*

Из определений, представленных в классической литературе по теории информации, следует: информация I и энтропия H обладают безразмерностью, а присутствие линейности их связи – одинаковостью их единиц. Поэтому далее будем говорить об энтропии, поскольку её важным аспектом является присутствие перед логарифмом вероятности (см. (2) и (3)). Она используется либо для усреднения информации, либо указывает на предположение о вероятном появлении события i , наделенного частной информацией: $I_i = \log_2 p_i$. Рассматривая дискретные случайные события, за неопределенность поведения i -го элемента примем энтропию, равную единице:

$$H(p) = -(p_i \log_2 p_i + q_i \log_2 q_i) = 1. \quad (8)$$

Представив (8) как

$$H^* = H^*(p_i) + H^*(q_i) = 1, \quad (9)$$

можно установить, что составляющие выражения (9) подлежат нормировке. Надстрочный символ (*) означает наличие в формулах нормированной величины энтропии.

В нашей задаче об определении количества информации нормировке будут принадлежать такие переменные, как частная информация, так и энтропия. При этом диапазон разброса этих переменных принимается от 0 до максимального значения.

Простейший вариант нормировки энтропии – линейное преобразование по формуле

$$H_i^* = \frac{H_i - H_{i\min}}{H_{i\max} - H_{i\min}}, \quad (10)$$

где $H_{i\min} = 0$. Все значения H^* по (10) будут укладываться в единичный отрезок $[0; 1]$.

Применение линейной нормировки можно считать приемлемым, когда значения переменной H^* плотно заполняют выделенный интервал. Если в статистических событиях данных присутствуют редкие вы-

бросы, выделяющиеся на фоне типичного разброса данных, то эти выбросы установят масштаб нормировки, при котором основная масса значений нормированной переменной сосредоточится вблизи нуля: $H^* \ll 1$. В таких случаях рекомендуется ориентироваться при нормировке не на экстремальные значения, а на типичные, относящиеся к таким статистическим характеристикам, как среднее и дисперсия. В этом случае большинство данных будет иметь единичный масштаб, и типичные значения всех переменных будут сопоставимы.

Тем не менее, возьмем на вооружение линейную нормировку и покажем энтропийные характеристики $H^*(p)$ для независимых событий.

Пусть система состоит из двух элементов, для каждого из которых определены на основе статистических данных вероятности работоспособного p_i и неработоспособного q_i состояний с условием: $p_i + q_i = 1$. Сумма вероятностей разнообразия событий:

$$p_1 p_2 + p_1 q_2 + p_2 q_1 + q_1 q_2 = 1. \quad (11)$$

Это выражение является основанием для построения функции определения суммарной энтропии двух независимых элементов:

$$\begin{aligned} H_{\Sigma} = & p_1 p_2 \log_2 p_2 + p_1 p_2 \log_2 p_1 + \\ & + p_1 q_2 \log_2 q_2 + p_1 q_2 \log_2 p_1 + p_2 q_1 \log_2 q_1 + \\ & + p_2 q_1 \log_2 p_2 + q_1 q_2 \log_2 q_2 + q_1 q_2 \log_2 q_1. \end{aligned} \quad (12)$$

Выражение (12) перепишем в виде:

$$\begin{aligned} H_{\Sigma} = & p_1 H(p_2) + p_2 H(p_1) + p_1 H(q_2) + \\ & + q_2 H(p_1) + p_2 H(q_1) + q_1 H(p_2) + q_1 H(q_2) + \\ & + q_2 H(q_1) = H(p_1) + \\ & + H(q_1) + H(p_2) + H(q_2). \end{aligned} \quad (13)$$

В (12) вероятности «перемешаны», то есть имеются произведения вероятностей, относящихся к противоположным событиям. Поэтому применение (13) неудобно для получения, например энтропий, относящихся к разнообразным событиям. Однако можно воспользоваться принципом определения искомым вероятностей событий через вероятности, относящиеся только к одному из противоположных событий. Так, на основе теорем сложения и умножения вероятностей, выражение (11) можно переписать в следующем виде:

$$p_1 p_2 + q_1 + q_2 - q_1 q_2 = 1. \quad (14)$$

Представление (14) в такой форме (с выделением вероятностей событий, рассчитываемых отдельно по каждой из составляющих: $p_1 p_2$; $q_1 + q_2 - q_1 q_2$), наталкивает на мысль о преобразовании выражения (13), которое было бы удобным для расчетов совместных и условных энтропий.

Учитывая вышеизложенные предпосылки, применив нормировку с условиями $H^*(p_i) + H^*(q_i) = 1$ и $p_i + q_i = 1$, выражение (13), с учетом преобразований запишем в виде

$$\begin{aligned} H_{\Sigma}^* = & p_1 H^*(p_2) + p_2 H^*(p_1) + p_1 H^*(q_2) + q_2 H^*(p_1) + \\ & + p_2 H^*(q_1) + q_1 H^*(p_2) + q_1 H^*(q_2) + q_2 H^*(q_1) = \dots \\ = & p_1 H^*(p_2) + p_2 H^*(p_1) + H^*(q_1) + H^*(q_2) + \\ & + q_1 H^*(p_2) + q_2 H^*(p_1) = p_1 H^*(p_2) + p_2 H^*(p_1) + \\ & + H^*(q_1) + H^*(q_2) + q_1 [1 - H^*(q_2)] + q_2 [1 - H^*(q_1)]. \end{aligned} \quad (15)$$

В (15) первые два слагаемых не подлежали преобразованиям, поскольку они позволяют определить совместную энтропию двух событий с вероятностями p_1 и p_2 (см. работы [3, 4]). Остальные четыре слагаемых относятся к вычислению условной энтропии. Выражение для определения условной энтропии можно получить иным путем:

$$\begin{aligned} H^*(p_1) + H^*(p_2) + H^*(q_1) + H^*(q_2) - p_2 H^*(p_1) - p_1 H^*(p_2) = \dots \\ = H^*(q_1) + H^*(q_2) + q_1 H^*(p_2) + q_2 H^*(p_1) = \\ = H^*(q_1) + H^*(q_2) + q_1 [1 - H^*(q_2)] + q_2 [1 - H^*(q_1)]. \end{aligned} \quad (16)$$

В (15) и (16) слагаемые, позволяющие рассчитывать величины совместной и условной энтропий, выделены в отдельные группы по качественному признаку. Тем самым, выражения для определения нормированных значений будут выглядеть следующим образом:

– для совместной энтропии

$$H^*(p_1 p_2) = p_2 H^*(p_1) + p_1 H^*(p_2); \quad (17)$$

– для условной энтропии

$$H^*(q/p) = H^*(q_1) + H^*(q_2) + q_1[1 - H^*(q_2)] + q_2[1 - H^*(q_1)]. \quad (18)$$

Получив нормированные значения энтропии по (17) и (18), можно перейти к определению искомых значений каждой из составляющей энтропии в битах. Например, для составляющих совместной энтропии:

$$H(p_1) = \frac{H(p_1) \cdot p_2 H^*(p_1)}{H^*(p_1)},$$

$$H(p_2) = \frac{H(p_2) \cdot p_1 H^*(p_2)}{H^*(p_2)}, \quad (19)$$

для совместной энтропии

$$H(p_1 p_2) = H(p_1) + H(p_2). \quad (20)$$

Заключение

Рассматривая в процессе разработки проектов и эксплуатации техническую систему, востребованы выражения для определения меры неопределенности информации её структурного содержания. Если природа состояний элементов системы такова, что их вероятности противоположны, то заимствование классических энтропийных характеристик резко усложняет поиск значений энтропии. Расширив свойства информации и введя условие нормировки в построение математических выражений, ориентированных на раздельный учет противоположных событий, решение задачи определения энтропии будет достижимым.

Полученные выражения (17)–(20) (на примере двух элементов) позволяют найти количество совместной энтропии для пересекающихся событий и условной энтропии для всех остальных противоположных событий.

Работа выполнена при поддержке гранта РФФИ № 15-08-01473.

Список литературы

1. Дулесов А.С., Агеева П.А. Мера информации в задаче оценки бесперебойной работы технической системы // Журнал «Фундаментальные исследования». – 2011. – № 12 (часть 1). – С. 102–107.

2. Дулесов А.С., Кабаева Е.В. Логарифмическая мера информации состояния технического объекта // Журнал «Современные проблемы науки и образования». – 2013. – № 1 (Электронный журнал); URL: www.science-education.ru/107-8210 (дата обращения: 28.01.2013).

3. Дулесов А.С., Кондрат Н.Н. Количество информации при наложении и пересечении элементарных событий // Журнал «Научное обозрение». – 2014. – № 12. – С. 146–150.

4. Дулесов А.С., Кондрат Н.Н. Мера неопределенности информации и её свойства применительно к оценке случайного поведения технического объекта // Журнал «Научное обозрение». – 2014. – № 7. – С. 258–264.

5. Дулесов А.С., Семенова М.Ю., Хрусталева В.И. Свойства энтропии технической системы // Журнал «Фундаментальные исследования». – 2011. – № 8 (часть 3). – С. 631–636.

6. Дулесов А.С., Ускова Е.А. Применение подходов Хартли и Шеннона к задачам определения количества информации технических систем // Журнал «Вопросы современной науки и практики. Университет им. В.И. Вернадского». – 2009. – № 2(16). – С. 46–49.

7. Дулесов А.С., Ускова Е.А. Применение формулы Хартли для оценки структурных связей элементов в задаче обеспечения надежного функционирования технических систем // Журнал «Вопросы современной науки и практики. Университет им. В.И. Вернадского». – 2009. – № 6(20). – С. 73–77.

8. Саночкин В.В. Что такое информация // Журнал «Эволюция». – 2005. – № 2. – С. 110–113.

9. Чернавский Д.С. Синергетика и информация (динамическая теория информации) / Изд. 2-е, испр. и доп. – М.: Едиториал УРСС. – 2004. – 288 с.

References

1. Dulesov A.S., Ageeva P.A. Mera informacii v zadache ocenki besperebojnoj raboty tehnicheckoj sistemy // Zhurnal «Fundamentalnye issledovanija». 2011. no. 12 (chast 1). pp. 102–107.

2. Dulesov A.S., Kabaeva E.V. Logarifmicheskaja mera informacii sostojanija tehnicheckogo obekta // Zhurnal «Sovremennye problemy nauki i obrazovanija». 2013. no. 1 (Elektronnyj zhurnal); URL: www.science-education.ru/107-8210 (data obrashhenija: 28.01.2013).

3. Dulesov A.S., Kondrat N.N. Kolichestvo informacii pri nalozhenii i peresechenii jelementarnyh sobytij // Zhurnal «Nauchnoe obozrenie». 2014. no. 12. pp. 146–150.

4. Dulesov A.S., Kondrat N.N. Mera neopredelennosti informacii i ejo svojstva primenitelno k ocenke sluchajnogo povedenija tehnicheckogo obekta // Zhurnal «Nauchnoe obozrenie». 2014. no. 7. pp. 258–264.

5. Dulesov A.S., Semenova M.Ju., Hrustaleva V.I. Svojtva jentropii tehnicheckoj sistemy // Zhurnal «Fundamentalnye issledovanija». 2011. no. 8 (chast 3). pp. 631–636.

6. Dulesov A.S., Uskova E.A. Primenenie podhodov Hartli i Shennona k zadacham opredelenija kolichestva informacii tehnicheckih sistem // Zhurnal «Voprosy sovremennoj nauki i praktiki. Universitet im. V.I. Vernadskogo». 2009. no. 2(16). pp. 46–49.

7. Dulesov A.S., Uskova E.A. Primenenie formuly Hartli dlja ocenki strukturnyh svjazej jelementov v zadache obespechenija nadezhnogo funkcionirovanija tehnicheckih sistem // Zhurnal «Voprosy sovremennoj nauki i praktiki. Universitet im. V.I. Vernadskogo». 2009. no. 6(20). pp. 73–77.

8. Sanochkin V.V. Chto takoe informacija // Zhurnal «Jevoljucija». 2005. no. 2. pp. 110–113.

9. Chernavskij D.S. Sinergetika i informacija (dinamicheskaja teorija informacii) / Izd. 2-e, ispr. i dop. M.: Editorial URSS. 2004. 288 p.

Рецензенты:

Гафнер Ю.Я., д.ф.-м.н., зав. кафедрой общей и экспериментальной физики, ХГУ им. Н.Ф. Катанова, г. Абакан;

Кочетков В.П., д.т.н., профессор, профессор кафедры электроэнергетики Хакасского технического института – филиала Сибирского федерального университета, г. Абакан.