

DOI [10.28925/2663-4023.2021.11.136143](https://doi.org/10.28925/2663-4023.2021.11.136143)

УДК. 004.056

Захарченко Микола Васильович

Доктор технічних наук, професор, професор кафедри
Кафедра кібербезпеки та технічного захисту інформації
Одеська національна академія зв'язку ім. О.С. Попова, Одеса
ORCID ID: 0000-0001-8946-7798
kaf.ibpd@onat.edu.ua

Гаджиев Матін Магсуд-огли

Доктор технічних наук, професор, професор кафедри
Кафедра кібербезпеки та технічного захисту інформації
Одеська національна академія зв'язку ім. О.С. Попова, Одеса
ORCID ID: 0000-0001-7280-3863
gadjevmm@ukr.net

Салманов Наріман Сулейман-огли

Кандидат технічних наук
Начальник «Телеком Сумгаїт» м. Сумгаїт, Азербайджан
ORCID ID: 0000-0002-0247-2517
nariman_s@box.az

Швець Наталія Василівна

Старший викладач
Кафедра інформаційних технологій та кібербезпеки
Одеська національна академія харчових технологій, Одеса
ORCID ID: 0000-0002-6719-3842
shvetsnv0601@gmail.com

Гавель Сергій Миколайович

Викладач
Кафедра кібербезпеки та технічного захисту інформації
Одеська національна академія зв'язку ім. О.С. Попова, Одеса
ORCID ID: 0000-0002-0484-5620
arkominer@gmail.com

АНАЛІЗ І ОЦІНКА ЯКІСНИХ І КІЛЬКІСНИХ ПОКАЗНИКІВ ІНФОРМАЦІЇ ПРИ ВИРІШЕННІ ЗАВДАНЬ ПОБУДОВИ СИСТЕМ ПЕРЕДАЧІ І ПЕРЕТВОРЕННЯ ДАНИХ

Анотація. Пояснення поняття інформації лише трохи доповнює інтуїтивне уявлення про цей термін і нічого не дає для побудови кількісної теорії інформації, яка могла б використовуватися при вирішенні інженерних задач. Для такої теорії необхідно ввести «операційне» визначення цього поняття, в основі якого лежить вказівка на спосіб вимірювання. Для оцінки кількості одержуваної інформації необхідно визначити міру невизначеності тій чи іншій ситуації. При передачі інформації на відстань на вході передавального перетворювача є кінцева система сигналів $[X, p]$, однозначно відповідних символам $\{A\}$, які утворюють *вхідний алфавіт каналу зв'язку*. На виході приймального пристрою існує кінцева система сигналів $[Z, p]$, що утворюють *вихідний алфавіт каналу зв'язку*. Процес передачі інформації полягає у виборі елементарних сигналів $x_k \in \{X\}$, посиленні їх по лінії зв'язку і винесення за прийнятим елементарним сигналом $z_l \in \{Z\}$ рішення по переданому сигналу x_k . Очевидно, що вилучення з сигналів $z_l \in \{Z\}$ інформації про сигнали $x_k \in \{X\}$ пов'язано з поданням алфавітів $\{X\}$ і $\{Z\}$ у вигляді об'єднаної множини $\{XZ\}$. Передача інформації в системах зв'язку завжди пов'язана зі



зміною у одержувача відомостей про фактично передані сигнали. Алфавіт джерела з точки зору одержувача має невизначеність, оскільки йому можуть бути відомі лише апіорні ймовірності P_k , але невідомо, чи був сигнал X_k переданий фактично; про це можна судити лише по величині апостеріорних ймовірностей. Цю невизначеність можна розглядати як міру бракуючої інформації, тобто як кількість інформації, необхідну для усунення цієї невизначеності. З цієї точки зору завдання одних числових характеристик випадкової величини недостатньо. У даній статті аналізуються питання достовірності подій в «ансамблі» прийнятих повідомлень в залежності від апіорних ймовірностей подій. Визначена оцінка кількості інформації в співвідношенні з достовірністю прийнятих повідомлень, позначені рекомендації, що використовуються при вирішенні більшості завдань, пов'язаних з побудовою систем передачі і перетворення інформації.

Ключові слова: інформація; кількість інформації; якість інформації; міра інформації; «ансамбль»; ймовірність подій; ентропія.

*Кожен хоче, щоб його інформували чесно, неупереджено,
правдиво - і в повній відповідності з його поглядами
Гілберт Честертон*

ВСТУП

Поняття інформації (її кількість і якість) відноситься до числа найважливіших понять в сучасній науці. З цим поняттям фахівці зустрічаються при вивченні будь-якого явища чи події. Аналіз історичного розвитку інформаційних процесів виявляє дві характерні особливості:

- порівняно донедавна отримана інформація носила в основному лише змістовний характер, і використання її для цілей управління було вкрай недостатньо [1];
- розвиток засобів передачі інформації і сам процес її передачі протікали повільно.

Поняття "інформація" може бути витлумачено як деяка сукупність відомостей, що визначають міру наших знань про ті чи інші події, явища або факти [2]. Насправді змістовний характер інформації може бути надзвичайно різноманітним: але у всіх випадках, в кінцевому підсумку реалізація процесу передачі, прийому і перетворення даних виконується із застосуванням загальновідомих класичних методів обробки і передачі даних або їх різновидів [3]. Відповідно різна інформація буде викликати різні емоції і буде представляти різну цінність.

Основні співвідношення, що визначають кількісну міру інформації, та основні теореми теорій інформації були сформульовані К. Шенноном і опубліковані ще в 1949 році [1]. Піонерами в розвитку теорії інформації були видатні вчені А. Колмогоров, О. Хінчин, О. Харкевич, Р. Добродушін, М. Железнов і інші. [2]. Ними була розвинена теорія кількісної міри інформації при рівноможливих подіях.

Слід звернути увагу на ту обставину, що будь-яка інформація з'являється при проведенні будь-якого дослідження, дії [1]. Чим можна охарактеризувати факт отримання інформації? Можна сказати, що факт отримання інформації пов'язаний з проведенням певного дослідження (дії) [4]. До дослідження має місце більша або менша невизначеність в вивчаємій ситуації або в результаті тих чи інших подій. Після дослідження (дії), тобто отримання інформації, ситуація стає більш визначеною і на поставлене запитання можна відповісти або однозначно, або, принаймні, кількість можливих різних відповідей стає менше, тобто невизначеність зменшується [1]. Таким чином, після отримання інформації хоча і залишається деяка невизначеність, але вона стає меншою.

В якості міри кількості інформації, отриманої в результаті того чи іншого дослідження (дії), можна взяти значення як функцію відносин [2]

$$K = \frac{n}{n_c}, \quad (1)$$

де n – кількість рівноможливих відповідей до дослідження, n_c – кількість відповідей після дослідження.

Знаючи числа n і n_c можна визначити ймовірності до і після дослідження [6]

$$\left. \begin{aligned} P_i(x_i) &= \frac{1}{n} \\ P(x_c) &= \frac{1}{n_c} \\ \frac{n}{n_c} &= \frac{P(x_i)}{P(x_c)} \end{aligned} \right\} \quad (2),$$

де $P_i(x_i)$ – ймовірність події до дослідження, $P(x_c)$ – ймовірність події після дослідження, $\frac{n}{n_c}$ – відношення ймовірностей.

Найбільш зручною для оцінювання стану подій (дослідження) є логарифмічна функція [6]

$$I(i) = \log_2 \frac{P_c(x_i)}{P(x_i)} \text{ дв. од.} \quad (3),$$

Тобто кількість інформації при двійковому логарифмі оцінюється в двійкових одиницях, які називаються бітами. Таким чином, отримання однієї двійкової одиниці кількості інформації відповідає тому, що ми дізнаємося, яка з двох рівноможливих подій має місце, або яка з двох рівноможливих гіпотез правильна. Такими двома подіями (гіпотезами) можуть бути відповіді "Так" або "Ні" на будь-яке питання. Якщо ці події рівноймовірні, то отримавши одну з них, ми тим самим отримуємо одну двійкову одиницю інформації [2].

Якщо події (відповіді) є рівноможливими і якщо до того ж $P_c(i) = 1$, тобто після дослідження ситуація повністю визначена, формула (3) може бути представлена у вигляді:

$$I(i) = \log_2 n \quad (4),$$

Така міра кількості інформації була запропонована в 1926р Р. Хартлі [1]. З викладеного та подальших міркуваннях видно, що формула (4) відноситься лише до вельми окремого випадку. Р. Хартлі правильно поставив питання про необхідність вимірювання кількості інформації [1]. Остаточна ця задача була вирішена в 1949 р К. Шенноном.

Розглянемо деяку кінцеву множину X подій, апріорні ймовірності яких дорівнюють [5]

$$P(x_1), P(x_2), P(x_3), \dots, P(x_n) \quad (5)$$

при умові, що:

$$P(x_1) + P(x_2) + P(x_3) \dots P(x_n) = 1$$

Останнє означає, що в перебігу деякого спостережуваного відрізка часу завжди відбувається одна з цих подій. Множину з відомим розподілом ймовірностей його елементів прийнято називати «ансамблем».

Розглянутий «ансамбль» подій може бути описаний кінцевою схемою (6)

$$x = \left(\frac{x_1 x_2 x_3 \dots x_n}{P(x_1) P(x_2) P(x_3) \dots P(x_n)} \right) \quad (6).$$

Використовуючи формулу (3), можна сказати, що достовірне повідомлення про те, що з усіх подій x відбувається подія x_i , несе в собі кількість інформації, рівне:

$$I(x_i) = \log \frac{1}{P(x_i)} \text{ дв. од.} \quad (7).$$

Отже, повідомлення про події несе тим більше інформації, чим менше його апріорна ймовірність [6]. Формула (7) показує, що в кінцевому ансамблі x повідомлення про різні події в загальному випадку несуть різну кількість інформації. При вирішенні багатьох завдань, пов'язаних з побудовою систем передачі і перетворення інформації, використовується середня кількість інформації, яка називається ентропією повідомлення "Н" [6].

$$\left. \begin{aligned} H(x) &= M\{I(x_i)\} \\ H(x) &= -\sum_{i=1}^n P(x_i) \log P(x_i) \end{aligned} \right\} \quad (8).$$

Як приклад, визначимо $H(x)$ при подіях:

$$P(x_1) = P \quad \text{та} \quad P(x_2) = 1 - P$$

У цьому випадку:

$$H(x) = -P \log P - (1 - p) \log(1 - p) \text{ дв. од.} \quad (9)$$

На Рис. 1 наведено залежність ентропії H від величини P :

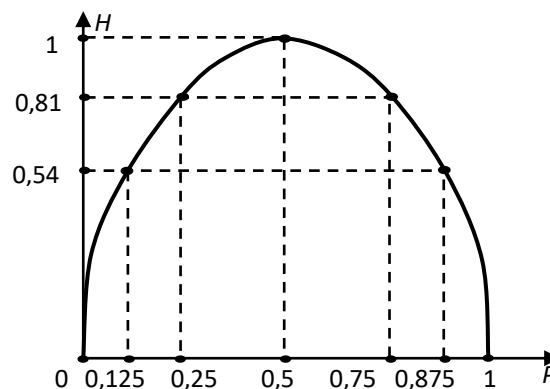


Рис. 1. Залежність ентропії від ймовірності

З рис.1 випливає:

- а) максимальне значення ентропії дорівнює при $P = 0,5$ ($H = 1$);
- б) при $P = 1$ і $P = 0$, $H = 0$.

Відзначимо основні властивості ентропії:

1. $H = 0$ лише в тому випадку, коли всі ймовірності $P(x_i) = 0$, [6] крім однієї, і ця єдина ймовірність дорівнює 1. Отже, $H = 0$ тільки в разі повної визначеності результату дослідження, а в решті випадків $H > 0$.

2. При заданому "n" ентропія максимальна і дорівнює

$$H(x) = \log n$$

$$P(x_1) = P(x_2) = P(x_3) = \dots = P(x_n) = \frac{1}{n} \quad (10).$$

Інтуїтивно зрозуміло, що такий стан відповідає найбільшій невизначеності. Дана властивість ентропії при $n = 2$ зображена на рис. 1.

3. При наявності двох ансамблів X і Y описуються схемами:

$$\left. \begin{aligned} X &= \frac{x_1, x_2, x_3, \dots, x_n}{P(x_1), P(x_2), P(x_3), \dots, P(x_n)} \\ Y &= \frac{y_1, y_2, y_3, \dots, y_n}{P(y_1), P(y_2), P(y_3), \dots, P(y_n)} \end{aligned} \right\} \quad (11)$$

При об'єднанні двох ансамблів схема об'єднання має вигляд, наведений в таблиці 1.

Таблиця 1

Схема об'єднання двох ансамблів

x_i Y_i	x_1	x_2	x_3	x_4	...	x_n
y_1	$P(x_1 y_1)$...				$P(x_n y_1)$
y_2						
·						
·						
·						
y_n	$P(x_1 y_n)$ $P(x_1)$...				$P(x_n y_n)$

Кількість інформації, що отримується в середньому на один дослід, може бути знайдена як зміна невизначеності в результаті досліді і отже

$$I(x, y) = I(x) - I(x/y) \quad (12)$$

Отже, взаємна інформація, що доставляється символом y про символ x дорівнює різниці інформацій, необхідних для однозначного визначення символу x до і після прийому символу y .

Взаємна інформація $I(x; y)$ може бути також представлена в формі

$$I(x; y) = I(x) + I(y) - I(x y) \quad (13),$$

де власна інформація пари символів $(x y)$ становить

$$I(x y) = -\log_2 p(x y) \quad (14)$$

З (13) випливає, що

$$I(x\ y) = I(x) + I(y) - I(x\ y), \quad (15),$$

тобто власна інформація пари символів $(x\ y)$ дорівнює власної інформації цих символів, обчисленої в припущенні їх незалежності, мінус взаємна інформація цих символів.

При дослідженні каналів зв'язку з декількома входами і виходами, наприклад, при рознесеній передачі або при рознесеному прийомі, виникає задача оцінки взаємної інформації n символів. Для двох символів x_1 і x_2 взаємна інформація визначається виразом (13).

ВИСНОВКИ:

1 Теорія зв'язку на сучасному етапі її розвитку не цікавиться семантикою, а використовуються терміни "інформація", "передача" і "перетворення даних", сприйняття яких засновані на властивості матерії - відображенні.

2 У даній роботі проведено дослідження, аналіз і оцінку деяких кількісних і якісних показників інформаційних процесів.

3 Показано, що при дослідженні каналів зв'язку з декількома входами і виходами для оцінки взаємної інформації двох символів x_1 і x_2 можна скористатися формулою $I(x; y) = I(x) + I(y) - I(x\ y)$.

4 Виконано передумови для оцінки взаємної інформації n символів, що має суттєве значення при вирішенні більшості завдань, пов'язаних з побудовою систем передачі і перетворення інформації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- 1 Зюко, А., Фалько, А., Панфилов, И., & Банкет, В. (1985). *Помехоустойчивость и эффективность систем передачи информации*. Радиосвязь.
- 2 Захарченко, М.В. (2009). *Системы передавання даних том 1. Заводостійке кодування*. Фенікс.
- 3 Цымбал, В.П. (1976). *Задачник по теории информации и кодированию*. «Вища школа».
- 4 Korchinskyi, V., Hadzhyiev, M., Pozdniakov, P., Kildishev, V., & Hordiichuk, V. (2018). Development of the procedure for forming nonstationary signal structures based on multicomponent LFM signals. *Eastern-European Journal of Enterprise Technologies*, 6(9 (96)), 29–38. <https://doi.org/10.15587/1729-4061.2018.151816>
- 5 Buriachok, V., Hadzhyiev, M., Sokolov, V., Skladannyi, P., & Kuzmenko, L. (2019). Implantation of indexing optimization technology for highly specialized terms based on Metaphone phonetical algorithm. *Eastern-European Journal of Enterprise Technologies*, 5(2 (101)), 43–50. <https://doi.org/10.15587/1729-4061.2019.181943>
- 6 Zaharchenko, M., Hadzhyiev, M., Salmanov, N., Golev, D., & Shvets, N. (2020). INFORMATION PARAMETERS OF CODES THAT ARE SYNTHESIZED ON THE BASIS OF ONE MODULE. *Cybersecurity: Education, Science, Technique*, 3(7), 95–102. <https://doi.org/10.28925/2663-4023.2020.7.95102>



Zaharchenko Mikola

Doctor of Technical Sciences, Professor,
Professor of Department of Cybersecurity and Technical Information Protection
Odessa National Academy of Telecommunications O. S. Popova
ORCID ID: 0000-0001-8946-7798
kaf.ibpd@onat.edu.ua

Hadzhyev Matin

Doctor of Technical Sciences, Professor
Professor of Department of Cybersecurity and Technical Information Protection
Odessa National Academy of Telecommunications O.S. Popova
ORCID ID: 0000-0001-7280-3863
gadjievmm@ukr.net

Salmanov Nariman

Ph.D
Head of department «Telekom Sumgait», Sumgait, Azerbaijan
ORCID ID: 0000-0002-0247-2517
nariman_s@box.az

Shvets Natalya

Senior Lecturer
Department of Information Technologies and Cybersecurity
Odessa National Academy of Food Technologies
ORCID ID: 0000-0002-6719-3842
shvetsnv0601@gmail.com

Havel Sergey Nikolaevich

Lecturer
Department of Cybersecurity and Technical Information Protection
Odessa National Academy of Telecommunications O.S. Popova
ORCID ID 0000-0002-0484-5620
arkominer@gmail.com

**ANALYSIS AND EVALUATION OF QUALITATIVE AND
QUANTITATIVE INDICATORS OF INFORMATION WHEN SOLVING
PROBLEMS OF CONSTRUCTION SYSTEMS OF DATA
TRANSMISSION AND TRANSFORMATION**

Abstract. The clarification of the concept of information only slightly supplements the intuitive understanding of this term and does not provide anything for the construction of a quantitative theory of information that could be used in solving engineering problems. For such a theory, it is necessary to introduce an "operational" definition of this concept, which is based on an indication of the measurement method. To estimate the amount of information received, it is necessary to find a measure of the uncertainty of a particular situation. When transmitting information over a distance, at the output of the transmitting converter there is a finite system of signals that unambiguously correspond to the symbols and form the input alphabet of the communication channel. At the output of the receiving device, there is a finite system of signals that form the output alphabet of the communication channel. The process of transmitting information consists in selecting chips, sending them over the communication line and making a decision on the transmitted signal based on the received chip. It is obvious that the extraction of information about signals from signals is associated with the representation of the alphabets $\{X\}$ and $\{Z\}$ in the form of a combined set $\{XZ\}$. The transfer of information in communication systems is always associated with a change at the receiver of information about the actually transmitted signals. From



the point of view of the receiver, the alphabet of the source has uncertainty, since he can only know a priori probabilities, but it is not known whether the signal was actually transmitted; this can be judged only by the magnitude of the posterior probabilities. From this point of view, setting some numerical characteristics of a random variable is not enough. This article analyzes the issues of reliability of events in the "ensemble" of received messages, depending on the a priori probabilities of events. The assessment of the amount of information in relation to the reliability of the received messages is made, the recommendations used in solving majority of the problems associated with the construction of information transmission and transformation systems are indicated.

Keywords. information; quantity of information, qualities of information, information measure , "ensemble", probability of events, entropy.

REFERENCES:

- 1 Zyuko, A., Falko, A., Panfilov, I., & Banquet, V. (1985). Noise immunity and efficiency of information transmission systems. Radio communication.
- 2 Zakharchenko, M.V. (2009). Data transmission systems volume 1. Factory production. Fenix.
- 3 Tsymbal, V.P. (1976). Problem book on information theory and coding. "Vishcha School".
- 4 Korchynskiy, V., Hadzhyiev, M., Pozdniakov, P., Kildishev, V., & Hordiichuk, V. (2018). Development of the procedure for forming nonstationary signal structures based on multicomponent LFM signals. *Eastern-European Journal of Enterprise Technologies*, 6(9 (96)), 29–38. <https://doi.org/10.15587/1729-4061.2018.151816>
- 5 Buriachok, V., Hadzhyiev, M., Sokolov, V., Skladannyi, P., & Kuzmenko, L. (2019). Implantation of indexing optimization technology for highly specialized terms based on Metaphone phonetical algorithm. *Eastern-European Journal of Enterprise Technologies*, 5(2 (101)), 43–50. <https://doi.org/10.15587/1729-4061.2019.181943>
- 6 Zaharchenko, M., Hadzhyiev, M., Salmanov, N., Golev, D., & Shvets, N. (2020). INFORMATION PARAMETERS OF CODES THAT ARE SYNTHESIZED ON THE BASIS OF ONE MODULE. *Cybersecurity: Education, Science, Technique*, 3(7), 95–102. <https://doi.org/10.28925/2663-4023.2020.7.95102>



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.