

УДК 004.056:519.1

А.О. Авсентьев

Определение ценности информации

Рассмотрен подход к определению численного значения обобщенного показателя ценности информации на основе логико-лингвистической модели оценки свойств, характеризующих этот показатель, и метода попарных сравнений, используемого для определения численных значений коэффициентов важности этих свойств.

Ключевые слова: информация, свойства информации, обобщенный показатель ценности информации, коэффициент важности, метод попарных сравнений.

doi: 10.21293/1818-0442-2016-19-1-21-24

Информация как обеспечивающий ресурс какой-либо деятельности характеризуется ее основными свойствами. Количество этих свойств и их возможные состояния обуславливают ценность информации для обеспечиваемой деятельности [1–3].

В процессе информационного обмена с использованием различных каналов и линий связи осуществляются различного рода преобразования этой информации, обусловленные условиями реализации процедур ее обработки, хранения и изменения материального носителя [1]. При этом соответствующие изменения претерпевают и свойства информации.

В [1] рассматривается подход к описанию рассматриваемых процедур как композиции операторов на множестве состояний этой информации при ее преобразовании из некоторого исходного состояния в конечное, т.е. в состояние, в котором информация будет принята потребителем. Кроме того, при формировании обобщенного показателя ценности информации в этой работе рассматриваются известные подходы к определению такого рода показателя. При этом считается, что каждое свойство информации измеримо [1–3].

Вследствие разнородности различных свойств информации способы измерения их численных значений также разнородны [4, 5]. По этой же причине, как показано в [1–5], степень важности рассматриваемых свойств для обеспечиваемой деятельности может быть определена только экспертным путем.

В настоящей работе предлагается решение сформулированных в [1] задач количественной оценки разнородных свойств информации и коэффициентов их важности с учетом потребностей конкретной деятельности. При этом для оценки численных значений рассматриваемых свойств информации предлагается использование аппарата теории нечетких множеств, а определение соответствующих коэффициентов важности осуществляется с помощью метода анализа иерархий Саати [6].

Формирование перечня свойств информации, характеризующего ее ценность для обеспечиваемой деятельности

Перечень свойств информации, характеризующий ее ценность, определяется видом и целями обеспечиваемой деятельности [1, 7]. Этим обстоятельством обусловлена тесная взаимосвязь разно-

родных свойств информации в рамках обобщенного показателя ее ценности.

По аналогии с [1] обозначим множество C свойств информации в виде выражения

$$C = \{c_i / i = 1, 2, \dots, I\},$$

где c_i – i -е свойство информации.

В качестве основных свойств, характеризующих ценность информации для обеспечиваемой деятельности, в [1, 7, 8] рассматриваются: достоверность, полнота, конфиденциальность, скрытность, своевременность, адекватность, релевантность, толерантность, понятность, разборчивость и др.

c_1 – достоверность – соответствие полученной информации действительной обстановке [1];

c_2 – полнота информации – её достаточность для принятия правильных решений [1, 7];

c_3 – конфиденциальность – это свойство, позволяющее не давать права на доступ к информации или не раскрывать ее содержание полномочным лицам [8];

c_4 – скрытность – это свойство, позволяющее скрыть факт передачи информации [1];

c_5 – своевременность – это поступление информации не позже заранее назначенного момента времени, согласованного со временем решения поставленной задачи [8];

c_6 – адекватность – уровень соответствия виду обеспечиваемой деятельности;

c_7 – релевантность – это показатель информации, который характеризует соответствие ее потребностям обеспечиваемой деятельности [7];

c_8 – толерантность – это показатель, характеризующий удобство восприятия и использования информации в процессе решения задач обеспечиваемой деятельности [7];

c_9 – понятность – это соответствие словарю языка общения, отражающему в полной мере лексику, используемую в обеспечиваемой деятельности;

c_{10} – разборчивость – это показатель, характеризующий правильность интерпретации информации пользователями и др.

Рассмотрим взаимосвязь перечисленных свойств в рамках деятельности конфиденциального характера, осуществляемой в условиях угроз безопасности информации.

Использование различного рода физических полей для передачи информации от источника потребителю в интересах обеспечиваемой деятельности по образованному для этих целей основному каналу связи [1–3] обуславливает формирование побочного канала связи, который используется злоумышленником для перехвата этой информации и является источником угроз ее безопасности.

Очевидно, что в этих условиях достоверная информация (свойство c_1), отражающая лишь некоторые аспекты обеспечиваемой деятельности, т.е. информация, имеющая недостаточную степень полноты (свойство c_2), не может обеспечить требуемую степень эффективности этой деятельности.

С другой стороны, достоверная и с достаточной степенью полноты отражающая основные аспекты обеспечиваемой деятельности информация в условиях нарушения ее конфиденциальности (свойство c_3) может быть использована злоумышленником для снижения эффективности этой деятельности. При этом при невозможности получения доступа к содержанию конфиденциальной информации для злоумышленника может представлять интерес сам факт информационного обмена (нарушение свойства c_4) в интересах обеспечиваемой деятельности с целью его использования для блокирования такого обмена или снижения показателя c_5 своевременности информации.

Своевременность информации характеризует временные характеристики информационного обмена. При этом достоверная, полная, конфиденциальная и скрытно переданная информация при ее преждевременном поступлении, когда потребитель не готов к ее восприятию, не может быть эффективно использована в интересах обеспечиваемой деятельности. С другой стороны, информация подвержена фактору старения, и превышение времени ее прохождения от источника к получателю некоторого значения, соответствующего этому фактору, может уже не соответствовать реальному состоянию этой деятельности.

При высоких значениях показателей $c_1 - c_5$ свойств информации и низком значении показателя c_6 ее адекватности информация не может быть использована в качестве обеспечивающего ресурса рассматриваемой деятельности.

Показатель c_7 свойства релевантности информации, характеризующий избыточность передаваемых в процессе информационного обмена сообщений, определяется как отношение объема полезной для обеспечиваемой деятельности части информации, содержащейся в этих сообщениях, к общему их объему. При этом увеличение общего объема сообщений может привести к соответствующему увеличению времени прохождения информации, т.е. к снижению показателя c_5 своевременности информации.

Низкое значение показателя c_8 толерантности информации, характеризующего удобство ее восприятия и использования в процессе решения задач обеспечиваемой деятельности, в свою очередь может

привести к увеличению времени обработки информации, передаваемой по каналам связи от источника к получателю, т.е. к снижению показателя c_5 ее своевременности.

Показатели c_9 и c_{10} характеризуют, по сути, возможность использования полученной информации в интересах обеспечиваемой деятельности. При этом их слишком низкое значение при приемлемых показателях $c_1 - c_4$ и $c_6 - c_8$ также может привести к снижению показателя c_5 своевременности информации, связанному с необходимостью повторной передачи непонятных или неразборчивых ее фрагментов.

Очевидно, что перечисленные свойства разнородны, и измерение их численных значений представляется задачей достаточно сложной. Однако следует отметить, что данные свойства объединяют два обстоятельства. Первое обстоятельство состоит в том, что данные свойства характеризуют ценность одной и той же информации или ее фрагмента – сообщения. Второе обстоятельство обусловлено использованием этой информации в качестве обеспечивающего ресурса некоторой деятельности.

Аналитически взаимосвязь перечисленных свойств информации применительно к некоторому сообщению x из их множества $x \in X$ по аналогии с [1] можно представить в виде обобщенного показателя ценности:

$$Z(x) = \sum_{i=1}^I w_i(x) c_i(x), \quad (1)$$

где $c_i(x)$ – набор измеримых показателей свойств информации для сообщения x , ($i = 1, \dots, 10$); $w_i(x)$ – весовые коэффициенты, характеризующие важность соответствующих свойств информации для рассматриваемой деятельности, определяемые экспертным путем; I – количество свойств информации, определяющих ценность сообщения.

Логико-лингвистическая модель для численной оценки свойств информации и их актуальности для формирования обобщенного показателя ее ценности

В связи с разнородностью указанных свойств информации для определения их численных значений целесообразно введение некоторой безразмерной величины $V_{\min} \leq V \leq V_{\max}$.

В качестве одного из способов решения данной задачи может быть использован лингвистический подход [9–12]. В рамках этого подхода каждый показатель, характеризующий свойство информации c_i , представляется некоторой лингвистической переменной V_i .

Формально лингвистические переменные определяются кортежем

$$\langle \beta, V, U \rangle, \quad (2)$$

где β – наименование лингвистической переменной; $V = \{V_1, \dots, V_i, \dots, V_I\}$ – базовое терм-множество, или множество значений лингвистических переменных, представляющих собой наименования нечетких переменных, областью определения которых является четкое множество U .

В таблице приведены значения элементов кортежа (2) для лингвистической переменной «Достоверность информации c_1 ».

Элементы кортежа для лингвистической переменной «Достоверность информации c_1 »

β	V_1	V_2	V_3	V_4	V_5	U
Достоверность информации, c_1	Очень низкая	Низкая	Средняя	Высокая	Очень высокая	0–100

Нечеткие переменные, в свою очередь, определяются кортежем [12]

$$\langle V, U, \hat{V}_i \rangle, \quad (3)$$

где V_i – наименование нечеткой переменной, $V_i \in \{V\}$; $U = \{U_i\}$ – область ее определения, или универсальное множество; $\hat{V}_i = \{\mu, U\}$ – нечеткое множество на U , описывающее ограничения на возможные числовые значения нечеткой переменной.

$\mu: U \Rightarrow [0, 1]$ – отображение множества U в единичный отрезок $[0, 1]$ – функция принадлежности нечеткого множества V .

На рис. 1 представлена графическая интерпретация для лингвистической переменной «Достоверность информации c_1 ».

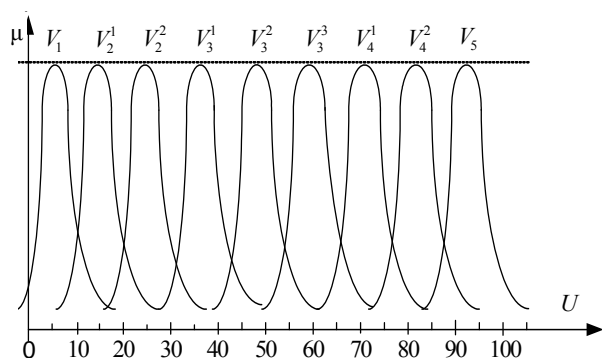


Рис. 1. Функции принадлежности лингвистической переменной «Достоверность информации c_1 »

Предложенная формализация нечетких понятий и отношений обеспечивается введением соответствующих нечетких и лингвистических переменных, нечеткого множества и их взаимосвязей [11, 12]. Это позволяет, с одной стороны, обеспечить переход от словесных описаний свойств информации к числовым, с другой стороны, дает возможность производить числовую оценку актуальности исследуемого свойства информации для обобщенного показателя ее ценности.

Множество значений нечетких переменных $\{V\}$ и области определения четкого множества $\{U\}$, а также вид функции принадлежности μ определяются экспертным путем применительно к виду обеспечиваемой деятельности.

Определение коэффициентов важности свойств информации методом Саати

С привлечением экспертов определим важность описанных выше свойств информации, обеспечи-

вающей деятельность конфиденциального характера [11]. Для этого расположим свойства c_i в порядке их приоритета для исследуемой деятельности:

$$c_3, c_1, c_2, c_5, c_6, c_4, c_7, c_8, c_9, c_{10}.$$

Определим лингвистические значения экспертных оценок приоритетных критериев коэффициентов важности $w_i(x)$ в (1):

- одинаковые;
- несколько больше;
- существенно больше;
- значительно больше;
- абсолютно больше;
- не сравнимы.

В соответствии с подходом, изложенным в [6], коэффициенты w_i должны удовлетворять условию

$$\sum_{i=1}^I w_i = 1.$$

Матрица $W = (w_{ij})$ попарных сравнений для этих коэффициентов строится с учетом следующего правила:

$$w_{ij} = \begin{cases} 1, & \text{если одинаково важные;} \\ 3, & \text{если несколько больше;} \\ 5, & \text{если существенно больше;} \\ 7, & \text{если значительно больше;} \\ 9, & \text{если абсолютно больше;} \\ 0, & \text{если не сравнимы.} \end{cases} \quad (4)$$

При условии $w_{ji} = (w_{ij})^{-1}$ матрица $W = (w_{ij})$ для свойств информации, описанных выше, с учетом критериев их приоритетных оценок будет иметь вид

$$W = \begin{pmatrix} 1 & 3 & 1 & 7 & 3 & 5 & 5 & 7 & 9 & 9 \\ 0,333 & 1 & 0,333 & 7 & 3 & 3 & 5 & 7 & 9 & 9 \\ 1 & 3 & 1 & 7 & 3 & 5 & 7 & 9 & 9 & 9 \\ 0,143 & 0,143 & 0,143 & 1 & 0,333 & 0,2 & 0,333 & 3 & 5 & 7 \\ 0,333 & 0,333 & 0,333 & 3 & 1 & 3 & 5 & 5 & 7 & 9 \\ 0,333 & 0,2 & 0,2 & 5 & 0,333 & 1 & 3 & 7 & 7 & 9 \\ 0,2 & 0,2 & 0,143 & 3 & 0,143 & 0,333 & 1 & 5 & 7 & 9 \\ 0,143 & 0,143 & 0,111 & 0,333 & 0,2 & 0,143 & 0,2 & 1 & 3 & 5 \\ 0,111 & 0,111 & 0,111 & 0,2 & 0,143 & 0,143 & 0,143 & 0,333 & 1 & 3 \\ 0,111 & 0,111 & 0,111 & 0,2 & 0,111 & 0,111 & 0,111 & 0,2 & 0,333 & 1 \end{pmatrix}$$

Вычисления коэффициентов w_i , выполненные с использованием программного продукта Mathcad 15, дают следующие результаты:

$$w_1 = 0,239; w_2 = 0,155; w_3 = 0,254; w_4 = 0,038; w_5 = 0,116; w_6 = 0,09; w_7 = 0,058; w_8 = 0,024; w_9 = 0,015; w_{10} = 0,011.$$

В качестве примера для c_i свойств информации, имеющих коэффициент важности $w_i > 0,1$, определим $c_i = 1; i = 1, 2, 3$, для свойств c_i с коэффициентом важности $0,1 > w_i \geq 0,05$ определим $c_i = 0,75; i = 5, 6, 7$, при $w_i < 0,05$ определим $c_i = 0,5; i = 4, 8, 9, 10$. Тогда значение $Z(x)$ обобщенного показателя ценности информации применительно к сообщению x в соответствии с (1) будет иметь вид

$$Z(x) = 0,239 \cdot 1 + 0,155 \cdot 1 + 0,254 \cdot 1 + 0,116 \cdot 0,75 + 0,09 \cdot 0,75 + 0,058 \cdot 0,75 + 0,038 \cdot 0,5 + 0,024 \cdot 0,5 + 0,015 \cdot 0,5 + 0,011 \cdot 0,5 = 0,89.$$

Очевидно, что максимальное значение обобщенного показателя ценности информации будет при одинаково важных свойствах c_i , т.е. при $w_i = 0,1$ и очень высоких значениях показателей всех

свойств, характеризующих информацию, $c_i = 1$, $Z(x)_{\max} = 1$.

Заключение

Таким образом, предложенный подход позволяет получить численное значение обобщенного показателя ценности информации как ресурса обеспечиваемой деятельности, а показатель (1) при этом может быть использован в качестве критерия для оценки характеристик каналов связи и различного рода преобразующих устройств, описанных в [1, 2].

Литература

1. Авсентьев А.О. Формирование обобщенного показателя ценности информации в каналах связи / А.О. Авсентьев, О.С. Авсентьев // Вестник Воронежского института МВД России. – 2015. – № 3. – С. 55–63.
2. Авсентьев А.О. Моделирование и оптимизация процессов передачи и защиты информации в каналах связи / А.О. Авсентьев, О.С. Авсентьев, В.В. Меньших // Специальная техника. – 2015. – № 5. – С. 47–50.
3. Авсентьев А.О. Модель оптимизации процесса обеспечения требований к свойствам информации при ее передаче по каналам связи / А.О. Авсентьев, В.В. Меньших // Вестник Воронежского института МВД России. – 2015. – № 4. – С. 147–154.
4. Шанкин Г.П. Ценность информации. Вопросы теории и приложений. – М.: Филоматис, 2004. – 128 с.
5. Morbhead D.R. The value of information and computer-aided information seeking: problem formulation and application to fiction retrieval / D.R. Morbhead, A.M. Pejtersen, W.B. Rouse // Information Processing & Management. – 1984. – Vol. 20, № 5–6. – P. 583–601.
6. Саати Т. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993. – 320 с.
7. Основы информационной безопасности: учебник для высших учебных заведений МВД России / под ред. В.А. Минаева и С.В. Скрыля. – Воронеж: Воронежский институт МВД России, 2001. – 464 с.

8. Парфенов В.И. Защита информации: словарь. – Воронеж: ГНИИИПТЗИ ФСТЭК России, 2003. – 210 с.

9. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976. – 168 с.

10. Модели принятия решений на основе лингвистической переменной / А.Н. Борисов и др. – Рига: Зинатне, 1982. – 256 с.

11. Катина А.М. Экспертная оценка системы знаний / А.М. Катина, С.В. Шидловский // Доклады ТУСУРа. – 2006. – № 5 (13). – С. 36–39.

12. Авсентьев О.С. Обеспечение конфиденциальности при подготовке специалистов по защите информации для органов внутренних дел: модели управления. – Воронеж: Воронежский институт МВД России, 2007. – 185 с.

Авсентьев Александр Олегович

Канд. техн. наук, преподаватель каф. физики
Воронежского института МВД России
Эл. почта: aoaao8787@mail.ru

Avsent'ev A.O.

Estimation of the information value

The article deals with the approach of the determination of numeral meaning of the general information index, based on the logical-linguistic model for assessment of properties, that characterize this index. Moreover, the method of pairwise comparisons, that is used to determine the numeral meaning of the importance coefficients of these properties is described.

Keywords: information, information properties, general information index, coefficient of importance, method of pairwise comparisons.