Определяне ценността на информацията в информационните системи

Костадин Костадинов

Estimating the Value of the Information in the Information System: *The method for an estimation of size value of the information about functioning information systems is offered. Connection between value of the information and means for its protection is defined.*

Key words. *Information, value, protection*

ВЪВЕДЕНИЕ

В теорията за активните информационни системи АС се приема, че информацията е единствен вид ресурс, който не се изтощава, а напротив при непрекъснат процес на информационна дифузия между елементите на системата, тя постоянно се увеличава. Освен това е известно, че информацията получена от интелектуалния приемник от външния свят, винаги се явява контекстно-независима и знанията получени чрез тях, чрез проведения анализ на възможностите на тяхното използване за достигане на поставената цел винаги се явяват контекстно-зависими от знанията натрупани в системата.Последното трябва да удовлетворява различни изисквания към качеството на информацията, най-важното от които се явява нейната „ценност”.

В многобройните работи, посветени на ценността на информацията, практически не се отчита характерната за участниците в АС целенасоченост (активност) на поведението или се изследва взаимодействието на активните информационни агенти с фиксирана роля, намиращи се на различни нива в йерархията [2,4]. При това се акцентира на факта, че информацията притежава два атрибута – избор на един вариант или няколко от множеството и втори – необходимата информация трябва да бъде съхранена или разпространена [3]. За целта информационната система трябва да бъде способна – да бъде измерима и да се съхранява, да се разпространява и да може да служи за анализ от други информационни източници. Системата още трябва да бъде и дисипативна [5].

Особено актуално, ценността на информацията в последните години се оказа масова информатизация на всички структури, когато комуникационните линии за връзка, бази и хранилища на информация, информационни, аналитически и експертни системи станаха необходими атрибути. Когато „ценността” в най-голяма степен помага за достигане на целите във функционирането на АС. Допълнително възникнаха за решение и особени задачи за защита корпоративната информация в държавните и регионални органи на властта.

1. Постановка на проблема. Нека е дадена информационна система Sis_{inf}^{n} функционираща в информационно пространства Sp_{inf}^{n} с размерност M_{inf}^{n} във време T_{sis}^{fun}. Да предположим, системата Sis_{inf}^{n} (в бъдеще – само “система”) се състои от n елемента, всеки от които се описва със скаларни параметри $X(t)$ за кой да е момент от време $t \in T_{sis}^{fun}$.

В практиката често възниква необходимост от определянето на основните категории задаващи величината на ценността на информацията $c(\xi)$ при функционирането на информационната система Sis_{inf}^{n} в пространството Sp_{inf}^{n}.

2. Същност на ценността на информацията. Информация Inf^{Vs} като вид ресурс трябва да удовлетворява различни изисквания, най-важното от които се явява нейната „ценност”. Понятието „ценност” на информацията $c(\xi)$, е неотменно

нейно качество и то се дефинира, като максимална полза, която се преследва от потребителя.

Определение 2.1. Под ценност $c(\xi)$ на информацията Inf^{Vs} се разбира величината на максимална полза, която дадено количество и качество информация Inf^{Vs} придобито в средата за функциониране на конкурентна система Sis_{inf}^{n} е способно да допринесе на потребителя с цел намаляване на „условните средни загуби" в адекватност, целенасоченост или друга мярка за нейното използване, поради флуктуацията на ентропията $H_{In_{ad}^{vs}}(\zeta)$ на Inf^{Vs} определена във време $t \in T_{sis}^{fun}$.

Ценността на информацията $c(\xi)$ [2.1.] се свързва с резултатите от информационният процес с понятието „условни средни загуби" Z_{inf}^{sr} в определена ситуация, които са независими (зависимите не се отчитат). Те могат да се опишат със зависимостта:

$$Z_{inf}^{sr} = \Delta Inf_{une}^{Vs} n_{бр}^{inf} \qquad (2.1.)$$

където - Z_{inf}^{sr} средни условни загуби на информация,%; ΔInf_{une}^{Vs} - стойност на информацията от една ситуация; $n_{бр}^{inf}$ - общ брой на ситуациите в информационно пространство Sp_{inf}^{n}.

Средните загуби Z_{inf}^{sr} по (2.1.) се определят по $n_{бр}^{inf}$ ситуации и тези - подбрани по релевантност, които са важни, характерни и ценни - $r_{бр}^{Inf^{Vs}}$. Тогава отношението $\frac{r_{бр}^{Inf^{Vs}}}{n_{бр}^{inf}} = P_{cenn}^{inf}$ формира вероятност на $c(\xi)$ - P_{cenn}^{inf}. При това за определяне на P_{cenn}^{inf} може да се използва плътността на бета-разпределение определено по закона на максималното правдоподобие[4].

Ценността на информацията $c(\xi)$ като категория, винаги преследва една цел – постоянно намаляване на общата ентропия $H_{In_{ad}^{vs}}(\zeta) \to min(\zeta)$. В този случай величина $c(\xi)$ има пряко отношение към понятието „ефективност" за работа на Sis_{inf}^{n} дефинирано с определени показатели. Те могат да притежават - положителна или отрицателна стойност. Например при $c(\xi) \le 0$ се приемат - дезинформация или нулева информация, при която P_{cenn}^{inf} постоянно намалява. При $c(\xi) = 0$ се формира нулева информация, която се формира от – параметри, събития и ситуации, които нямат никакво съотношение към определен информационен процес; повтаряща се информация в пределите от време не изискващо никаква реакция за промяна функционирането на Sis_{inf}^{n}; обща нефилтрирана за достоверност информация и др. Величината на $c(\xi)$ също се измерва с корелацията на други показатели на „ценност", за които системата функционира. Във всички случаи ценността се явява особена форма за минимизация на общата ентропията $H_{In_{ad}^{vs}}(\zeta) \to min(\zeta)$ в дискретни състояния ξ.

В тези състояния в системата се генерира неопределеност $H_{In_{ad}^{vs}}(c(\zeta))$, която може да се определи:

$$H_{In_{ad}^{vs}}(c(\zeta)) = -b \sum_{H_{In_{ad}^{vs}}(c(\zeta)) \in Sis_{inf}^{n}} p_F(c(\xi)) ln\, p_F(c(\xi)) \qquad (2.2.)$$

където - $-b$ - коефициент на пропорционалност,(0-0.5).

Ако $c(\xi)$ е функционално фиксирана в (2.2.) за даден тип система,тогава оптималното разпределение (апостериорното бета-разпределение - $f(\Lambda)$) на вероятността p_F за времето T_{sis}^{fun} може да дефинира $H_{In_{ad}^{vs}}(c(\zeta))$ или

$$-\frac{df(\Lambda)T_{sis}^{fun}}{dT_{sis}^{fun}} = H_{In_{ad}^{vs}}(c(\zeta)) \qquad (2.3.)$$

При което $$\left(\frac{dH_{In_{ad}^{vs}}(c(\zeta))}{dT_{sis}^{fun}}\right) = -\frac{db\,ln\,p_F}{dT_{sis}^{fun}}T_{sis}^{fun} \qquad (2.4.)$$

и решението на уравнението е:

$$\frac{d}{dT}\left(H_{In_{ad}^{vs}}(c(\zeta))\right) = \frac{1}{2}\left(\frac{db\,ln\,p_F T_{sis}^{fun^2}}{dT_{sis}^{fun}} + C\right) \qquad (2.5.)$$

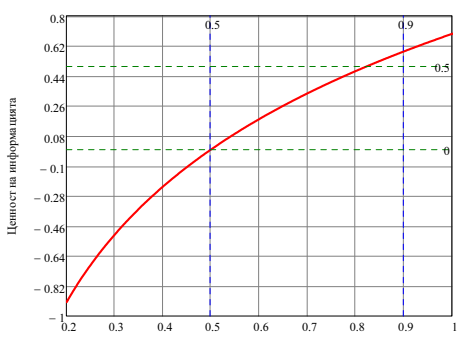
Функционалната неопределеност $H_{F_{ad}^{vs}}(c(\zeta))$ (2.3), заради ценността на информацията $c(\xi)$ вследствие притока на информация $dInf^{Vs}$ води до намаляване величина $H_{In_{ad}^{vs}}(c(\zeta)) - dI_{inf}^{sis} = H_{In_{ad}^{vs}}(c(\zeta))_{+inf\ f}^{sis}$, което е очевидно.

Ако приемем, че стойността на формулирана ценност $(c(\zeta))$ е предварително желана, то можем да изберем величините на вероятността $p_F = e^{\beta f(\Lambda)\in I_{\Omega_inf}^{Vs}(\zeta)}$ за осигуряване началното количество информация в системата. В този случай чрез въвеждане на допълнителна информация за обработка величината на $(c(\zeta))$ може да се определи от:

$$(c(\zeta)) = ln\left(\frac{p_{In_ihf+1}^{s}}{p_{In_inf\,0}^{s}}\right) \qquad (2.6.)$$

където $p_{ihf\,0}^{s}$ - вероятност за начално количество информация,(0-1); p_{inf+1}^{s} вероятност за нарастването на информацията получена от ситуациите релевантност,(0-1).

Определянето на $(c(\zeta))$ по (2.6.) при оптимални данни за функциониране на системата $p_{ihf\,0}^{s} = 0.5 - 0.6$ и $p_{In_inf+1}^{s} = 0.9$ след получаване на dI_{inf}^{sis} стойността на $(c(\zeta))$ е демонстрирана на Фиг.1.



Фиг.1. Ценност на информацията

Стойността на $(c(\zeta))$ при величина $p^s_{In_inf+1} \geq 0.5$ осигурява ценност $(c(\zeta)) > 0$ (Фиг.1.), а при $p^s_{In_inf+1} = 1$ съответно $(c(\zeta)) > 0.69315$. Следователно в практиката е необходимо да се говори за ценност на информацията тогава, когато е спазено условието за допълнителната информация $p^s_{In_inf+1}$ да надвишава началната, в стойностите $0.6 \leq p^s_{In_inf+1} \geq 0.99$.

Ценността на информацията $c(\xi)$ обхваща харлиевското, болцмановското или шенноновкото определение за количество информация Inf^{Vs} [5]. Показателите на същата могат да се изразят с временна, кумулативна и алгоритмично полезна стойност в зависимост от предпочитанията на потребителите.

3. Защита на информацията. Ценността на информацията Inf^{Vs} предполага действия по нейната защита от други информационни системи. При избора на средства за защита да се отчита правилото: „условната стойност" в сили и средства на защитата не трябва да превишава „условната стойност" изразходвана за реализиране ценността на информацията. За демонстрация на посоченото правило приемаме, че $c(\xi 1)$ е ценност на информацията придобита от собствената система $Sis^n_{inf}\,1$ (с мерки да я съхраним), а $c(\xi 2)$ е ценност на информацията на противоположната система - $Sis^n_{inf}\,2$ (мерки да добият информацията, ниво на изтичане на информация от собствената - $Sis^n_{inf}\,1$). За целта нека C_1^{zash}, C_2^{dob} - са средствата и действията, заделени съответно - за защита $Sis^n_{inf}\,1$ или добиване за $Sis^n_{inf}\,2$ на информация. Тогава - p_1^{zash}, p_2^{dob} - да бъдат вероятностите за успех на защитаваната информация в $Sis^n_{inf}\,1$ и съответно добиващата информация $Sis^n_{inf}\,2$.

Стойностите на p_1^{zash}, p_2^{dob} са трудни за математическо описание и последващо оценяване в определен момент от функционирането на системите във време $t \in T^{fun}_{sis}$, доколкото тя зависи от активността на $Sis^n_{inf}\,1$ и $Sis^n_{inf}\,2$ изразена чрез функция от $p = f\left(C_1^{zash}, C_2^{dob}\right)$. Знаменията на $p = f\left(C_1^{zash}, C_2^{dob}\right)$ предполагат гарантиране на успеха при функционирането и на двете системи при единно информационно пространство Sp^n_{inf}. При това – ситуациите изразени с вероятностите за „защитата" и „изтичането" информация за $Sis^n_{inf}\,1$ и $Sis^n_{inf}\,2$ съответно са несъвместими и се представят от пълна група ситуации, при което:

$$p_1^{zash} + p_2^{dob} = 1 \qquad (3.1.)$$

Тогава, вероятността определена от функцията $p = f\left(C_1^{zash}, C_2^{dob}\right)$ се определя по вероятностите на уравненията за:

защита на Inf_1^{Vs} $$p_1^{zash} = \frac{q_1^s.C_1^{zash}}{q_1^s.C_1^{zash} + q_2^s.C_2^{dob}} \qquad (3.2.)$$

изтичане Inf_2^{Vs} $$p_2^{dob} = \frac{q_2^s.C_2^{dob}}{q_2^s.C_2^{dob} + q_1^s.C_1^{zash}} \qquad (3.3.)$$

където q_1^s, q_2^s - е теглови коефициент и средствата заделени от конкуриращите системи $Sis^n_{inf}\,1$ и $Sis^n_{inf}\,2$ спрямо наличният им информационен потенциал за постигане на поставените цели.

Определение 3. Величината на защита на информацията C_1^{zash}, в даден тип система $Sis_{inf}^n 1$ в общо Sp_{inf}^n за функциониране се определя, ако се изразява потенциалната вероятност за изтичане на информация чрез неравенството:

$$C_1^{zash} \le \frac{q_1^s.C_1^{zash}}{q_1^s.C_1^{zash} + q_2^s.C_2^{dob}} c(\xi 1) \qquad (3.4.)$$

Отчитайки, че $C_1^{zash} \ge 0$ и $c(\xi 1) \ge 0$, ще получим системата неравенства:

$$q_1^s.\left(C_1^{zash}\right)^2 + q_2^s.C_1^{zash}C_2^{dob} - C_1^{zash}c(\xi 1) \le 0 \qquad (3.5.)$$

$$p_1^{zash}, p_2^{dob}, C_1^{zash}, C_2^{dob}, c(\xi 1) \ge 0 \qquad (3.6.)$$

Решавайки системи (3.5.) и (3.6.), величината на C_1^{zash} за защита на Inf_1^{Vs} в $Sis_{inf}^n 1$, ще състави:

$$C_1^{zash} = \frac{\sqrt{q_1^s.\left(C_1^{zash}\right)^2 + 4q_1^s, q_2^s C_1^{zash} c(\xi 1)} - q_1^s.C_1^{zash}}{2q_1^s} \qquad (3.7.)$$

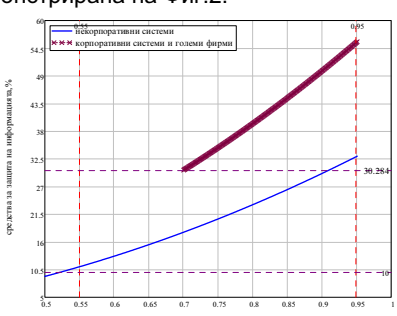
Ако предложим, че C_2^{dob} на страната използваща изтичащата информация е приблизително равна с тази $c(\xi 2)$, а информацията има еднаква ценност и за системи $Sis_{inf}^n 1$ и $Sis_{inf}^n 2$, то (3.7.) приема вида:

$$C_2^{dob} = c(\xi 2)\frac{\sqrt{q_2^s + 4q_1^s, q_2^s} - q_2^s}{2q_2^s} \qquad (3.8.)$$

Тогава ако конкуриращите системи $Sis_{inf}^n 1$ и $Sis_{inf}^n 2$ се намират в равни условия за изпълнение на своите задачи, т.е. $q_1^s = q_2^s$, то C_1^{zash} се измерва с C_2^{dob} за защита на информацията, при което C_2^{dob} се определя по приблизителната зависимост:

$$C_2^{dob} = c(\xi 2)\frac{\sqrt{5} - 1}{2} \qquad (3.9.)$$

Стойностите на (3.9.) зависят от типа на $Sis_{inf}^n 1$ и $Sis_{inf}^n 2$. За целта за осигуряване ценността на информацията за обикновените некорпоративни системи в нормите $0.5 \le c(\xi 2) \le 0.95$ разхода заделени средства за нейната защита е в пределите $10 \le C_1^{zash} \ge 33\%$. Математическата интерпретация осигуряване ценността на информацията е демонстрирана на Фиг.2.



Фиг. 2. Средства за защита ценността на информацията,%

Но при защита на ценността на информацията при корпоративни системи, научни центрове, големи фирми и др. осигуряващи ценност $0.7 \le c(\xi 2) \le 0.95$ разходите за защита на информацията са $30.284 \le C_1^{zash} \ge 55.778 \ \%$ Фиг.2.

ЗАКЛЮЧЕНИЕ

Предложен е метод за оценка ценностна на информацията при функционирането на информационните системи на основата отчитането средни условни загуби в информация. Изведена е зависимостта, че ценността на информацията се явява особена форма за минимизация на общата ентропията в информационните системи. Дефинирана е връзката между ценността на информацията, величините на информационната ефективност и величините на защита на информацията при конкуриращите системи, необходими за формулиране на решения за функциониране.

Литература

[1] Platt W., Strategic intelligence production, New Vork, 1957.
[2] Гоппа В.Д., Введение в алгебраическую теорию информации, М.: Наука. Физматлит, 1995.—112 с.
[3] Гасанов Э.Э., Теория сложности информационного поиска, МГУ имени М.В.Ломоносова, Москва 2005.
[4] Колганов С. К., Корников В. В., Попов П. Г., Хованов Н. В., Построение в условиях дефицита информации сводных оценок сложных систем, Москва, 1994.
[5] Крушный В.В., Основы теории информации и кодирования, Снежинск - СГФТА 2005.
[6] Соболь И.М., Метод Монте-Карло, М.: Наука, 1978.

За контакти:

Доц. д.т.н. Костадин Костадинов, Катедра “ОУТПСВ”, НВУ “Васил Левски”, тел.: 062-630 834, e-mail: koce_knk@abv.bg

Докладът е рецензиран.