



Оценка степени конфиденциальности документа на основе нечёткого анализа контекстных сведений

Информатика и автоматика

Сулейманова А.Н.

Бакинский государственный университет

E-mail: jafa25@mail.ru

Рассматриваются предпосылки, необходимые для своевременного установления адекватного уровня грифа конфиденциальности сведений, подлежащих включению в будущий документ. Для оценки факторов конфиденциальности применяются соответствующие экспертные заключения из удалённых источников, которые после предварительной обработки классифицируются с использованием системы нечёткого вывода. Присвоение документу грифа конфиденциальности осуществляется на основе агрегации экспертных оценок.

Ключевые слова: фактор конфиденциальности, экспертная оценка, нечёткое множество, система нечёткого вывода.

Введение

Создание условий для своевременного установления адекватного уровня грифа конфиденциальности сведений, подлежащих включению в будущий документ, является первым и, пожалуй, основным элементом защиты документированной информации, позволяющим обеспечить относительно надёжную безопасность служебной тайны в конкретной организации [1]. Понятно, что любая ценная информация, не закреплённая ни на одном носителе, находится в наибольшей безопасности [2]. Однако угрозы для её конфиденциальности появляются немедленно при возникновении мысли о необходимости её документирования. В связи с этим система защиты неподлежащей разглашению информации должна начинать функционировать не после её публикации (подписания), а заблаговременно, т.е. до момента нанесения на чистый носитель первых знаков будущего документа. Поэтому первоначально решаются вопросы: является ли данная информация конфиденциальной и, если да, то какой уровень грифа конфиденциальности ей должен быть присвоен?

В основе процесса присвоения грифа конфиденциальности сведений, подлежащих включению в будущий документ, лежат перечень конфиденциальных сведений организации, требования партнёров, условия контрактов, а также перечень конфиденциальных документов организации. При этом, система маркирования документов (присвоение им грифов) не способна полностью гарантировать сохранность конфиденциальной информации [3]. Тем не менее, она позволяет чётко регламентировать работу с документами, формируя в том числе систему доступа к ним персонала.

Постановка задачи

Предлагается использовать механизм нечёткого вывода для предварительной классификации и обработки экспертных заключений из удалённых источников относительно степени конфиденциальности сведений, подлежащих включению в будущий документ, и

применять их для последующей агрегации с целью выявления консолидированной оценки уровня секретности конкретного документа. Другими словами, предлагается осуществить нечёткий анализ факторов конфиденциальности (ФК) сведений с использованием экспертного опроса методом шкальных оценок. Более того, в процессе принятия решений относительно степени конфиденциальности сведений, подлежащих включению в оцениваемый документ, предварительные экспертные оценки степеней конфиденциальности рассматриваемых документов, как правило, выражаются всевозможными терминами, которые могут быть описаны в виде нечётких множеств, восстанавливаемых с помощью соответствующих функций принадлежности. Для упорядочения оцениваемых экспертами ФК, а значит и нечётких множеств, отражающих их степени предпочтительности, применяется метод нечёткого вывода.

Решение задачи

Степень конфиденциальности (DC) сведений, подлежащих включению в оцениваемый документ, традиционно определяется путём сопоставления выявленного экспертом уровня конфиденциальности сведений с заданным максимальным уровнем в системе применяемых критериев оценки конфиденциальности. Соответствующий показатель может быть вычислен по формуле [4]:

$$DC = \frac{\sum_{i=1}^K \alpha_i \frac{c_{ri}}{c_{di}}}{\max_i \sum_{i=1}^K \alpha_i \frac{c_{ri}}{c_{di}}} \times 100\%, \quad (1)$$

где K – число критериев оценки конфиденциальности; α_i – вес i -го критерия оценки конфиденциальности, определяющий степень его важности; c_{di} – максимальный уровень конфиденциальности сведений согласно i -му критерию оценки конфиденциальности; c_{ri} – выявленный экспертом уровень конфиденциальности сведений согласно i -му критерию оценки конфиденциальности. Определение весовых коэффициентов (весов) α_i и оценки отношения выявленного экспертом уровня конфиденциальности сведений с заданным максимальным уровнем в системе применяемых критериев оценки конфиденциальности $c_i = c_{ri}/c_{di}$, проводится с использованием экспертного опроса методом шкальных оценок. Обобщённый показатель консолидированного мнения всех экспертов относительно c_i должен удовлетворять стремлению [4]:

$$\sum_{i=1}^n w_i c_i \rightarrow \max, \quad (2)$$

при условии $\sum_{i=1}^n w_i = 1$, где w_i – значение весового коэффициента функции c_i . Тогда результирующее значение отношения выявленного экспертом уровня конфиденциальности сведений к заданному максимальному уровню в системе применяемых критериев оценки конфиденциальности (c_i) определяется в виде усреднения: $c_i = \frac{1}{m} \sum_{j=1}^m c_{ij}$, где m – число привлечённых экспертов; c_{ij} – оценка уровня конфиденциальности сведений со стороны j -го эксперта. При этом, степень согласованности (W) мнений групп экспертов в целом по совокупности всех ФК определяется как [4]:

$$W = \frac{12}{K^2 - K} \sum_{i=1}^K (c_i - \frac{K+1}{2})^2. \quad (3)$$

Обобщение мнений экспертов относительно значений α_i , определяющих относительную важность i -го критерия оценки конфиденциальности, осуществляется аналогично.

Оценка конфиденциальности сведений методом нечёткого логического вывода.

Пусть U является множеством альтернативных экспертных оценок степеней конфиденциальности сведений, а A – его нечётким подмножеством, принадлежность к которому элементов из U определяется соответствующими значениями из $[0; 1]$ так называемой *функции принадлежности* [5]. Предположим, что нечёткие множества A_j описывают возможные значения (термы) лингвистической переменной (ЛП) x [5]. Тогда множество решений (альтернатив) относительно уровня конфиденциальности рассматриваемого документа (КД) можно характеризовать совокупностью критериев – значениями ЛП: x_1, x_2, \dots, x_r . Например, значением «СУЩЕСТВЕННАЯ» ЛП $x_1 = \text{научная значимость}$. Совокупность термов

ЛП (или критериев), принимающих подобные значения, могут характеризовать представления о достаточности рассматриваемых экспертных оценок относительно степени конфиденциальности. Тогда, полагая $S=$ степень конфиденциальности также как ЛП, типовое правило может выглядеть как: «Если $x_1=$ СУЩЕСТВЕННАЯ и $x_2=$ НЕЗНАЧИТЕЛЬНАЯ, тогда $S=$ ВЫСОКАЯ».

В общем случае импликативные рассуждения руководителя службы КД (РКД) можно представить в следующем виде:

$$e_i: \langle \text{Если } x_1=A_{1i} \text{ и } x_2=A_{2i} \text{ и } \dots x_p=A_{pi}, \text{ то } S=B_i \rangle. \quad (4)$$

где A_{ki} ($k=1 \div p$) и B_i ($i=1, 2, \dots$) – нечёткие множества, отражающие термы входных и выходных ЛП соответственно. Для компьютерной реализации правил вида (4) нами применяется достаточно простая процедура фаззификации для термов из их левых частей правил (4), в соответствии с которой каждый такой терм отражается в виде нечёткого подмножества дискретного множества оцениваемых альтернатив $\{a_1, a_2, \dots, a_n\}$ в следующем виде:

$$A_k = \frac{\mu_{A_k}(a_1)}{a_1} + \frac{\mu_{A_k}(a_2)}{a_2} + \dots + \frac{\mu_{A_k}(a_n)}{a_n}, \quad (5)$$

где $\mu_{A_k}(a_t)$ ($t=1 \div n$) – значение функции принадлежности, восстанавливающее нечёткое множество A_k . В качестве таковой нами выбрана Гауссовская функция вида:

$$\mu_{A_k}(a_t) = \exp \left\{ -\frac{[e_k(a_t)-100]^2}{\sigma_k^2} \right\}, \quad (6)$$

где $e_k(a_t)$ – консолидированная экспертная оценка альтернативы a_t ($t=1 \div n$) по стобальной шкале на предмет соответствия сведений критерию оценки по k -му ФК; σ_k^2 – плотность распределения ближайших элементов, которую мы выбираем единой для всех случаев процесса фаззификации. Далее находим пересечение $A_i=A_{1i} \cap x_2=A_{2i} \cap \dots \cap x_p=A_{pi}$. В дискретном случае операция пересечения нечётких множеств определяется нахождением минимума соответствующих значений их функций принадлежности, т.е. путём нахождения [5]:

$$\mu_{A_i}(v) = \min\{\mu_{A_i}(u_1), \mu_{A_i}(u_2), \dots, \mu_{A_i}(u_p)\}, \quad (7)$$

где $V=U_1 \times U_2 \times \dots \times U_p$; $v=(u_1, u_2, \dots, u_p)$; $\mu_{A_i}(u_j)$ – степень принадлежности элемента u_j нечёткому множеству A_i . Тогда правила (4) можно представить в более компактном виде: d_i : «Если $x=A_i$, то $S=B_i$ ».

Для реализации нечётких импликативных правил используются различные операторы нечёткой импликации [5], например, импликация Лукасевича, которую в принятых обозначениях сформулируем как

$$\mu_H(w, i) = \min\{1, 1 - \mu_A(w) + \mu_B(i)\}, \quad (8)$$

где H – нечёткое подмножество на $W \times I$; $w \in W$ и $i \in I$. Аналогичным образом рассуждения (правила) e_1, e_2, \dots, e_p транспонируются в соответствующие нечёткие множества H_1, H_2, \dots, H_p . При этом, обозначая их произведение как $D=H_1 \cap H_2 \cap \dots \cap H_p$, для каждой пары $(w, i) \in W \times I$ согласно [5] получим: $\mu_D(w, i) = \min\{\mu_{H_j}(w, i)\}, j=1 \div p$. В этом случае вывод об удовлетворительности альтернативы, описанной нечётким подмножеством A из W , можно определить через композиционное правило $G=A \circ D$, где G является нечётким подмножеством единичного интервала I ; « \circ » обозначает операцию композиции правил, которая в принятых выше обозначениях выглядит как

$$\mu_G(i) = \max\{\min[\mu_A(u), \mu_D(w, i)]\}.$$

Сравнение альтернатив осуществляется на основе их точечных оценок. С этой целью в начале для нечёткого подмножества $C \subset I$ определяются α -уровневые множества ($\alpha \in [0; 1]$) в виде $C_\alpha = \{i | \mu_C(i) \geq \alpha, i \in I\}$. Затем для каждого из них определяются средние значения соответствующих элементов $M(C_\alpha)$. В общем случае для множества, состоящего из n элементов [6],

$$M(C_\alpha) = \sum_{j=1}^n \frac{i_j}{n}, i \in C_\alpha. \quad (9)$$

В итоге, численная оценка нечёткого множества C , отражающего степень удовлетворительности соответствующей альтернативы, получается из равенства [6]:

$$F(C) = \frac{1}{\alpha_{\max}} \int_0^{\alpha_{\max}} M(C_\alpha) d\alpha. \quad (10)$$

Классификация оценок влияния ФК на уровень секретности документа. Оценка степени конфиденциальности сведений, подлежащих включению в оцениваемый документ, являясь многокритериальной процедурой, подразумевает применение композиционного правила агрегирования оценки по каждому ФК сведений x_i ($i=1\div 6$). Поэтому для оценки такого влияния выберем пять оценочных понятий: u_1 – «НЕЗНАЧИТЕЛЬНОЕ»; u_2 – «ОЩУТИМОЕ»; u_3 – «СУЩЕСТВЕННОЕ»; u_4 – «ЗНАЧИТЕЛЬНОЕ»; u_5 – «ОЧЕНЬ БОЛЬШОЕ», характеризующих их степень влияния на уровень секретности документа (рисунок). Проще говоря, под множеством $C=(u_1, u_2, u_3, u_4, u_5)$ будем понимать совокупность признаков, по которым классифицируются значимости ФК. Тогда классификацию оценки влияния ФК на уровень секретности документа произведём с использованием достаточного набора нечётких импликативных правил вида: «Если $\langle \dots \rangle$, то $\langle \dots \rangle$ », и на их основе установим соответствующую шкалу градации будущих оценок.

Итак, для построения шкалы градации оценок выберем за основу следующие непротиворечивые рассуждения, которые могут и должны быть согласованы с РКД:

r_1 – «Если в документе присутствуют ФК x_1, x_2 и x_3 , то их совокупное влияние на уровень секретности существенное»;

r_2 – «Если вдобавок к указанным ФК имеет место и фактор x_6 , то в совокупности их влияние на уровень секретности будет более чем существенным»;

r_3 – «Если вдобавок к условиям, оговоренным в r_2 , имеет место фактор x_5 , то совокупное влияние всех ФК на уровень секретности будет очень существенным»;

r_4 – «Если в документе присутствуют все ФК, т.е. $x_1\div x_6$, то совокупное влияние их на уровень секретности будет чересчур существенным»;

r_5 – «Если же в документе присутствуют ФК x_1, x_2, x_3 и x_4 , а остальные: x_5 и x_6 – нет, то все равно совокупное влияние ФК на уровень секретности будет существенным»;

r_6 – «Если же в документе отсутствуют факторы x_1, x_2, x_4 и x_5 , то совокупное влияние ФК на уровень секретности будет несущественным».

В приведённых высказываниях (или причинно-следственных связях) входными характеристиками будем считать наличие (или отсутствие) ФК x_k ($k=1\div 6$), а выходной характеристикой y – степень их влияния на уровень секретности документа. Тогда, определив соответствующие значения (термы) лингвистических переменных x_k ($k=1\div 6$) и y , на базе приведённых высказываний построим нечёткие импликативные правила в виде:

r_1 : «Если x_1 =ИМЕЕТ МЕСТО и x_2 =ИМЕЕТ МЕСТО и x_3 =ИМЕЕТ МЕСТО, то y =СУЩЕСТВЕННОЕ»;

r_2 : «Если x_1 =ИМЕЕТ МЕСТО и x_2 =ИМЕЕТ МЕСТО и x_3 =ИМЕЕТ МЕСТО и x_6 =ИМЕЕТ МЕСТО, то y =БОЛЕЕ ЧЕМ СУЩЕСТВЕННОЕ»;

r_3 : «Если x_1 =ИМЕЕТ МЕСТО и x_2 =ИМЕЕТ МЕСТО и x_3 =ИМЕЕТ МЕСТО и x_5 =ИМЕЕТ МЕСТО и x_6 =ИМЕЕТ МЕСТО, то y =ОЧЕНЬ СУЩЕСТВЕННОЕ»;

r_4 : «Если x_1 =ИМЕЕТ МЕСТО и x_2 =ИМЕЕТ МЕСТО и x_3 =ИМЕЕТ МЕСТО и x_4 =ИМЕЕТ МЕСТО и x_5 =ИМЕЕТ МЕСТО и x_6 =ИМЕЕТ МЕСТО, то y =ЧЕРЕСЧУР СУЩЕСТВЕННОЕ»;

r_5 : «Если x_1 =ИМЕЕТ МЕСТО и x_2 =ИМЕЕТ МЕСТО и x_3 =ИМЕЕТ МЕСТО и x_4 =ИМЕЕТ МЕСТО и x_5 =НЕ ИМЕЕТ МЕСТО и x_6 =НЕ ИМЕЕТ МЕСТО, то y =СУЩЕСТВЕННОЕ»;

r_6 : «Если x_1 =НЕ ИМЕЕТ МЕСТО и x_2 =НЕ ИМЕЕТ МЕСТО и x_4 =НЕ ИМЕЕТ МЕСТО и x_5 =НЕ ИМЕЕТ МЕСТО, то y =НЕСУЩЕСТВЕННОЕ».

Выходную ЛП y зададим на дискретном множестве $J=\{0; 0.1; 0.2, \dots; 1\}$. Тогда $\forall j \in J$ используемые в правилах $r_1\div r_6$ её термы опишем нечёткими множествами с соответствующими функциями принадлежности [5]: Y_1 =СУЩЕСТВЕННОЕ: $\mu_{Y_1}(j)=j$; Y_2 =БОЛЕЕ ЧЕМ СУЩЕ-

СТВЕННОЕ: $\mu_{Y_2}(j)=j^{1/2}$; Y_3 =ОЧЕНЬ СУЩЕСТВЕННОЕ: $\mu_{Y_3}(j)=j^2$; Y_4 =ЧЕРЕСЧУР СУЩЕСТВЕННОЕ: $\mu_{Y_4}(j)=1$ при $j=1$ и $\mu_{Y_4}(j)=0$ при $j<1$; Y_0 =НЕСУЩЕСТВЕННОЕ: $\mu_{Y_0}(j)=1-j$.

Согласно подходу, описанному в [6], фаззификацию термов в левых частях правил осуществим с помощью Гауссовской функции принадлежности: $\mu(u_i)=\exp\{-(u_i-1)^2/\sigma_j^2\}$ ($j=1\div 5$), восстанавливающей нечёткие подмножества дискретного универсума (u_1, u_2, u_3, u_4, u_5), где $u_i=(d_{i-1}+d_i)/2$ ($i=1\div 5$). В данном случае значения для σ_j подбираются, исходя из степени важности ФК.

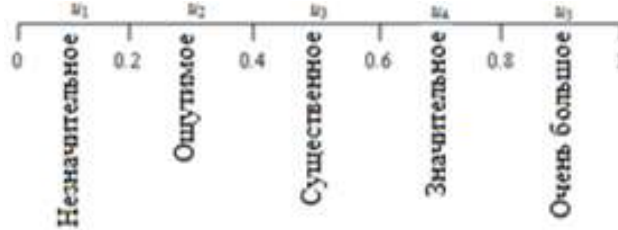


Рисунок. Градация влияния ФК в масштабе [0; 1]

Путём простого преобразования $x = d_0 + t(d_5 - d_0)$ ($t \in [0; 1]$) отрезок $[d_0; d_5]$ всегда можно легко трансформировать в отрезок $[0; 1]$ (рисунок). Поэтому, оценивая наличие ФК с точки зрения их значимости, градуированных в масштабе единичного интервала, где $d_i = -0.2i$ ($i=0\div 5$), все термы из левых частей правил $r_1\div r_6$ запишем в виде следующих нечётких множеств, полагая, что значимость изменений с возрастанием их порядкового номера растёт:

- ИМЕЕТ МЕСТО (ФК x_1): $A = \{0.0013/u_1; 0.0183/u_2; 0.1299/u_3; 0.4797/u_4; 0.9216/u_5\}$;
- ИМЕЕТ МЕСТО (ФК x_2): $B = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.3679/u_4; 0.8948/u_5\}$;
- ИМЕЕТ МЕСТО (ФК x_3): $C = \{0.0063/u_1; 0.0468/u_2; 0.2096/u_3; 0.5698/u_4; 0.9394/u_5\}$;
- ИМЕЕТ МЕСТО (ФК x_4): $D = \{0.0183/u_1; 0.0889/u_2; 0.2910/u_3; 0.6412/u_4; 0.9518/u_5\}$;
- ИМЕЕТ МЕСТО (ФК x_5): $E = \{0.0392/u_1; 0.1409/u_2; 0.3679/u_3; 0.6977/u_4; 0.9608/u_5\}$;
- ИМЕЕТ МЕСТО (ФК x_6): $F = \{0.0687/u_1; 0.1979/u_2; 0.4376/u_3; 0.7427/u_4; 0.9675/u_5\}$.

С учётом введённых формализмов правила $r_1\div r_6$ уже формулируются как:

- r_1 : «Если $x_1=A$ и $x_2=B$ и $x_3=C$, тогда $y=Y_1$ »;
- r_2 : «Если $x_1=A$ и $x_2=B$ и $x_3=C$ и $x_6=F$, тогда $y=Y_2$ »;
- r_3 : «Если $x_1=A$ и $x_2=B$ и $x_3=C$ и $x_5=E$ и $x_6=F$, тогда $y=Y_3$ »;
- r_4 : «Если $x_1=A$ и $x_2=B$ и $x_3=C$ и $x_4=D$ и $x_5=E$ и $x_6=F$, тогда $y=Y_4$ »;
- r_5 : «Если $x_1=A$ и $x_2=B$ и $x_3=C$ и $x_4=D$ и $x_5=\neg E$ и $x_6=\neg F$, тогда $y=Y_1$ »;
- r_6 : «Если $x_1=\neg A$ и $x_2=\neg B$ и $x_4=\neg D$ и $x_5=\neg E$, тогда $y=Y_0$ ».

Далее, для левых частей этих правил вычислим функции принадлежности $\mu_{M_i}(u)$ ($i=1\div 6$). В частности, согласно (7) имеем:

- r_1 : $\mu_{M_1} = \min\{\mu_A(u), \mu_B(u), \mu_C(u)\}$, $M_1 = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.3679/u_4; 0.8948/u_5\}$;
- r_2 : $\mu_{M_2} = \min\{\mu_A(u), \mu_B(u), \mu_C(u), \mu_F(u)\}$, $M_2 = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.3679/u_4; 0.8948/u_5\}$;
- r_3 : $\mu_{M_3} = \min\{\mu_A(u), \mu_B(u), \mu_C(u), \mu_E(u), \mu_F(u)\}$, $M_3 = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.3679/u_4; 0.8948/u_5\}$;
- r_4 : $\mu_{M_4} = \min\{\mu_A(u), \mu_B(u), \mu_C(u), \mu_D(u), \mu_E(u), \mu_F(u)\}$, $M_4 = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.3679/u_4; 0.8948/u_5\}$;
- r_5 : $\mu_{M_5} = \min\{\mu_A(u), \mu_B(u), \mu_C(u), \mu_D(u), 1-\mu_E(u), 1-\mu_F(u)\}$, $M_5 = \{0.0001/u_1; 0.0043/u_2; 0.0622/u_3; 0.2573/u_4; 0.0325/u_5\}$;
- r_6 : $\mu_{M_6} = \min\{1-\mu_A(u), 1-\mu_B(u), 1-\mu_D(u), 1-\mu_E(u)\}$, $M_6 = \{0.9608/u_1; 0.8591/u_2; 0.6321/u_3; 0.3023/u_4; 0.0392/u_5\}$.

В итоге правила запишутся в ещё более компактной форме:

- r_1 : «Если $x=M_1$, то $Y=Y_1$ »;
- r_2 : «Если $x=M_2$, то $Y=Y_2$ »;
- r_3 : «Если $x=M_3$, то $Y=Y_3$ »;
- r_4 : «Если $x=M_4$, то $Y=Y_4$ »;
- r_5 : «Если $x=M_5$, то $Y=Y_1$ »;
- r_6 : «Если $x=M_6$, то $Y=Y_0$ ».

Преобразования этих правил с помощью импликации (8) позволили получить нечёткие отношения в виде матриц R_1, R_2, \dots, R_6 , пересечение которых по средствам применения операции \min по всем соответствующим элементам в итоге дало следующее общее функциональное решение:

$$R = \begin{bmatrix} & 0 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1 \\ u_1 & 0.9999 & 0.9999 & 0.9992 & 0.9492 & 0.8792 & 0.7892 & 0.6792 & 0.5492 & 0.3992 & 0.2292 & 0.0392 \\ u_2 & 0.9957 & 0.9957 & 0.9957 & 0.9957 & 0.9809 & 0.8909 & 0.7809 & 0.6509 & 0.5009 & 0.3309 & 0.1409 \\ u_3 & 0.9378 & 0.9378 & 0.9378 & 0.9378 & 0.9378 & 0.9378 & 0.9378 & 0.8779 & 0.7279 & 0.5579 & 0.3679 \\ u_4 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6321 & 0.6977 \\ u_5 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.1052 & 0.9608 \end{bmatrix}$$

Согласно вышеприведённым рассуждениям, нечёткий вывод относительно k -го уровня ($k=1 \div 5$) наличия конфиденциальности в зависимости от соответствующего ФК отражается в виде нечёткого подмножества E_k дискретного конечного универсума J с соответствующими значениями функции принадлежности из k -ой строки матрицы R . В частности, вывод относительно наличия ФК 1-го уровня интерпретируется в виде нечёткого множества $E_1 = \{0.9999/0; 0.9999/0.1; 0.9999/0.2; 0.9492/0.3; 0.8792/0.4; 0.7892/0.5; 0.6792/0.6; 0.5492/0.7; 0.3992/0.8; 0.2292/0.9; 0.0392/1\}$. Устанавливая α -уровневые множества $E_{1\alpha}$ и вычисляя их мощности $M(E_{1\alpha})$ по формуле (9), имеем:

- для $0 < \alpha < 0.0392$: $\Delta\alpha = 0.0392$, $E_{1\alpha} = \{0; 0.1; 0.2; \dots, 0.9; 1\}$, $M(E_{1\alpha}) = 0.5$;
- для $0.0392 < \alpha < 0.2292$: $\Delta\alpha = 0.19$, $E_{1\alpha} = \{0; 0.1; 0.2; \dots, 0.8; 0.9\}$, $M(E_{1\alpha}) = 0.45$;
- для $0.2292 < \alpha < 0.3992$: $\Delta\alpha = 0.17$, $E_{1\alpha} = \{0; 0.1; 0.2; \dots, 0.7, 0.8\}$, $M(E_{1\alpha}) = 0.40$;
- для $0.3992 < \alpha < 0.5492$: $\Delta\alpha = 0.15$, $E_{1\alpha} = \{0; 0.1; 0.2; 0.3, 0.4, 0.5, 0.6, 0.7\}$, $M(E_{1\alpha}) = 0.35$;
- для $0.5492 < \alpha < 0.6792$: $\Delta\alpha = 0.13$, $E_{1\alpha} = \{0; 0.1; 0.2; 0.3, 0.4, 0.5, 0.6\}$, $M(E_{1\alpha}) = 0.30$;
- для $0.6792 < \alpha < 0.7892$: $\Delta\alpha = 0.11$, $E_{1\alpha} = \{0; 0.1; 0.2; 0.3, 0.4, 0.5\}$, $M(E_{1\alpha}) = 0.25$;
- для $0.7892 < \alpha < 0.8792$: $\Delta\alpha = 0.09$, $E_{1\alpha} = \{0; 0.1; 0.2; 0.3, 0.4\}$, $M(E_{1\alpha}) = 0.20$;
- для $0.8792 < \alpha < 0.9492$: $\Delta\alpha = 0.07$, $E_{1\alpha} = \{0; 0.1; 0.2; 0.3\}$, $M(E_{1\alpha}) = 0.15$;
- для $0.9492 < \alpha < 0.9992$: $\Delta\alpha = 0.05$, $E_{1\alpha} = \{0; 0.1; 0.2\}$, $M(E_{1\alpha}) = 0.10$;
- для $0.9992 < \alpha < 0.9999$: $\Delta\alpha = 0.0007$, $E_{1\alpha} = \{0; 0.1\}$, $M(E_{1\alpha}) = 0.05$.

В итоге, на основании (10) точечную оценку нечёткого выхода E_1 получим в виде

$$F(E_1) = \frac{1}{0.9999} (0.5 \cdot 0.0392 + 0.45 \cdot 0.19 + 0.4 \cdot 0.17 + 0.35 \cdot 0.15 + 0.3 \cdot 0.13 + 0.25 \cdot 0.11 + 0.2 \cdot 0.09 + 0.15 \cdot 0.07 + 0.1 \cdot 0.05 + 0.05 \cdot 0.0007) = 0.3257.$$

Аналогичными действиями устанавливаем точечные оценки для остальных выходов: при уровне оценочного понятия $u_2 - F(E_2) = 0.3647$; $u_3 - F(E_3) = 0.4350$; $u_4 - F(E_4) = 0.5470$; $u_5 - F(E_5) = 0.9453$. В данном случае, значение $F(E_1) = 0.3257$, являющееся дефазсифицированным выходом применённой нечёткой модели для комплексной оценки влияния ФК на уровень секретности сведений в документе, является верхней границей интервала, внутри которого эта оценка характеризуется как «ИМЕЕТ НЕСУЩЕСТВЕННОЕ ВЛИЯНИЕ». Аналогичным образом, выход: $F(E_2) = 0.3647$ является верхней границей для оценки «ИМЕЕТ СУЩЕСТВЕННОЕ ВЛИЯНИЕ»; $F(E_3) = 0.4350$ является верхней границей для оценки «ИМЕЕТ БОЛЕЕ ЧЕМ СУЩЕСТВЕННОЕ ВЛИЯНИЕ»; $F(E_4) = 0.5470$ является верхней границей для оценки «ИМЕЕТ ОЧЕНЬ СУЩЕСТВЕННОЕ ВЛИЯНИЕ»; $F(E_5) = 0.9453$ является верхней границей для оценки «ИМЕЕТ ЧЕРЕСЧУР СУЩЕСТВЕННОЕ ВЛИЯНИЕ». Тогда, выбрав в качестве критерия формирования итоговой оценки равенство: $E = \frac{F(E_k)}{F_{\max}} \times 100$, $k=1 \div 5$, где $F(E_k)$ – оценка k -го уровня наличия и влияния ФК (в широком смысле и любая другая оценка), $F_{\max} = F(E_5) = 0.9453$, в принятых допущениях получим обоснованную шкалу для оценки уровня конфиденциальности документа в масштабе отрезка $[0; 100]$, которая сведена в табл.1.

Таблица 1. Градация уровней конфиденциальности методом нечёткого вывода

Интервал	Уровень конфиденциальности
(57.86; 100]	«Строго конфиденциально»
(46.01; 57.86]	«Конфиденциально: под особым контролем»
(38.58; 46.01]	«Конфиденциально»
(34.45; 38.58]	«Для служебного пользования»
[0.00; 34.45]	«Не нуждается в грифе конфиденциальности»

Теперь представим, что экспертному сообществу предлагается по пятибалльной системе протестировать 10 альтернативных документов: a_k ($k=1\div 10$) на предмет оценки степени влияния ФК x_i ($i=1\div 6$) на уровень их секретности. Так, на основе консолидированных (усреднённых) экспертных заключений, удовлетворяющих требованиям (2) и согласованных согласно равенству (3), и применения к ним критерия итоговой оценки (1), для документов a_k ($k=1\div 10$) получены взвешенные оценки уровней конфиденциальности, которые сведены в табл.2.

Таблица 2. Классификация документов a_k по уровням конфиденциальности

№	Весовой коэффициент оценки ФК						Соотношение итоговой оценки (%)	Рекомендуемый гриф конфиденциальности для документа
	α_1	α_2	α_3	α_4	α_5	α_6		
	0.261	0.207	0.177	0.158	0.121	0.074		
	1	1	2	7	6	3		
a_1	4.67	4.67	4.80	4.53	4.47	4.67	93.27	«Строго конфиденциально»
a_2	3.87	4.47	4.47	4.20	4.27	4.47	85.19	«Конфиденциально: под особым контролем»
a_3	3.47	3.80	4.13	4.07	4.07	4.13	77.22	«Конфиденциально»
a_4	2.80	2.80	2.87	3.00	3.40	3.00	58.27	«Для служебного пользования»
a_5	2.00	2.40	2.73	2.27	2.73	2.47	47.90	«Не нуждается в грифе конфиденциальности»
a_6	4.47	4.60	4.67	4.73	4.47	4.53	91.62	«Строго конфиденциально»
a_7	3.87	4.20	4.07	4.60	4.00	4.27	81.91	«Конфиденциально: под особым контролем»
a_8	1.67	1.87	2.07	2.07	1.87	2.07	37.78	«Не нуждается в грифе конфиденциальности»
a_9	3.20	3.60	3.13	3.80	3.60	3.60	68.09	«Конфиденциально»
a_{10}	2.53	2.33	2.67	2.47	2.60	2.80	50.84	«Для служебного пользования»

Чтобы построить систему нечёткого логического вывода относительно оценок уровней конфиденциальности сведений в документах a_k ($k=1\div 10$), прошедших первичную экспертную экспертизу по пятибалльной системе оценивания на предмет влияния ФК x_i ($i=1\div 6$) (табл.2), воспользуемся конструкцией вербальной модели, сформулированной в Разд. 4 посредством высказываний $r_1\div r_6$. В этом случае для термов из их левых частей правил вида $r_1\div r_6$ процедура фаззификации применена несколько иначе, а именно, каждый терм отражён в виде нечёткого подмножества конечной совокупности оцениваемых альтернатив (в нашем случае документов) $\{a_1, a_2, \dots, a_{10}\}$ в следующем виде: $A_i = \{\mu_{A_i}(a_1)/a_1; \mu_{A_i}(a_2)/a_2; \dots; \mu_{A_i}(a_{10})/a_{10}\}$, где $\mu_{A_i}(a_k)$ ($k=1\div 10$) – значение функции принадлежности нечёткого множества A_i , устанавливающего насколько документ a_k соответствует критерию оценки A_i . В качестве таковой нами выбрана Гауссовская функция принадлежности $\mu_{A_i}(a_k) = \exp\{-[e_i(a_k)-5]^2/\sigma_i^2\}$, где $e_i(a_k)$ – усреднённая экспертная оценка по пятибалльной шкале относительно степени влияния i -го ФК на уровень секретности документа a_k ($k=1\div 10$); σ_i^2 – плотность расположения ближайших элементов, которую мы выбираем единой для всех случаев процесса фаззификации как равной 4. Тогда, полагая уровни вли-

яния ФК x_i ($i=1\div 6$) лингвистическими переменными, их максимально критичные термы, в частности, «ОЧЕНЬ БОЛЬШОЕ» представим в виде соответствующих нечётких подмножеств A_i дискретного универсума $U=\{a_1, a_2, \dots, a_{10}\}$ следующим образом:

- $A_1=\{0.9731/a_1; 0.7267/a_2; 0.5570/a_3; 0.2982/a_4; 0.1054/a_5; 0.9322/a_6; 0.7267/a_7; 0.0625/a_8; 0.4449/a_9; 0.2176/a_{10}\}$;
- $A_2=\{0.9731/a_1; 0.9322/a_2; 0.6977/a_3; 0.2982/a_4; 0.1845/a_5; 0.9608/a_6; 0.8521/a_7; 0.0864/a_8; 0.6126/a_9; 0.1683/a_{10}\}$;
- $A_3=\{0.9900/a_1; 0.9322/a_2; 0.8276/a_3; 0.3217/a_4; 0.2758/a_5; 0.9731/a_6; 0.8056/a_7; 0.1169/a_8; 0.4172/a_9; 0.2574/a_{10}\}$;
- $A_4=\{0.9463/a_1; 0.8521/a_2; 0.8056/a_3; 0.3679/a_4; 0.1552/a_5; 0.9819/a_6; 0.9608/a_7; 0.1169/a_8; 0.6977/a_9; 0.2019/a_{10}\}$;
- $A_5=\{0.9322/a_1; 0.8753/a_2; 0.8056/a_3; 0.5273/a_4; 0.2758/a_5; 0.9322/a_6; 0.7788/a_7; 0.0864/a_8; 0.6126/a_9; 0.2369/a_{10}\}$;
- $A_6=\{0.9731/a_1; 0.9322/a_2; 0.8276/a_3; 0.3679/a_4; 0.2019/a_5; 0.9463/a_6; 0.8753/a_7; 0.1169/a_8; 0.6126/a_9; 0.2982/a_{10}\}$.

Тогда с учётом этих формализмов и представленных в Разд. 5 формальных описаний термов из правых частей правил, базовая модель запишется в следующем виде:

- r_1 : «Если $x_1=A_1$ и $x_2=A_2$ и $x_3=A_3$, то $y=Y_1$ »;
- r_2 : «Если $x_1=A_1$ и $x_2=A_2$ и $x_3=A_3$ и $x_6=A_6$, то $y=Y_2$ »;
- r_3 : «Если $x_1=A_1$ и $x_2=A_2$ и $x_3=A_3$ и $x_5=A_5$ и $x_6=A_6$, то $y=Y_3$ »;
- r_4 : «Если $x_1=A_1$ и $x_2=A_2$ и $x_3=A_3$ и $x_4=A_4$ и $x_5=A_5$ и $x_6=A_6$, то $y=Y_4$ »;
- r_5 : «Если $x_1=A_1$ и $x_2=A_2$ и $x_3=A_3$ и $x_4=A_4$ и $x_5=\neg A_5$ и $x_6=\neg A_6$, то $y=Y_1$ »;
- r_6 : «Если $x_1=\neg A_1$ и $x_2=\neg A_2$ и $x_4=\neg A_4$ и $x_5=\neg A_5$, то $y=Y_0$ ».

Преобразование этих правил уже в привычном виде привело к функциональному решению \mathbf{R} , отражающему на конечном дискретном множестве $J=\{0; 0.1; \dots; 1\}$ причинно-следственную связь между консолидированными экспертными оценками документов по ФК x_i ($i=1\div 6$) и соответствующими уровнями их конфиденциальности. Согласно вышесуперупомянутым рассуждениям, k -я строка матрицы \mathbf{R} является нечётким выводом (множеством) относительно агрегированного уровня конфиденциальности для k -ой альтернативы (документа).

$$\mathbf{R} = \begin{bmatrix} & 0 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1 \\ a_1 & 0.0269 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.9731 \\ a_2 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.9322 \\ a_3 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.4430 & 0.8056 \\ a_4 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.7018 & 0.6273 & 0.5273 \\ a_5 & 0.8946 & 0.8946 & 0.8946 & 0.8946 & 0.8758 & 0.7758 & 0.6758 & 0.5758 & 0.4758 & 0.3758 & 0.2758 \\ a_6 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.0678 & 0.9819 \\ a_7 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.2733 & 0.9608 \\ a_8 & 0.9378 & 0.9378 & 0.9169 & 0.8169 & 0.7169 & 0.6169 & 0.5169 & 0.4169 & 0.3169 & 0.2169 & 0.1169 \\ a_9 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.5828 & 0.6977 \\ a_{10} & 0.8317 & 0.8317 & 0.8317 & 0.8317 & 0.8317 & 0.7369 & 0.6369 & 0.5369 & 0.4369 & 0.3369 & 0.2369 \end{bmatrix}$$

Чтобы численно интерпретировать каждый из нечётких выводов применим описанную выше процедуру дефаззификации. Так, для нечёткого вывода E_1 относительно уровня конфиденциальности 1-го документа, т.е. для $E_1=\{0.0269/0; 0.0678/0.1; 0.0678/0.2; 0.0678/0.3; 0.0678/0.4; 0.0678/0.5; 0.0678/0.6; 0.0678/0.7; 0.0678/0.8; 0.0678/0.9; 0.0678/1\}$, с учётом (9) имеем:

- для $0 < \alpha < 0.0269$: $\Delta\alpha=0.0269$; $E_{1\alpha}=\{0; 0.1; 0.2; \dots; 0.9; 1\}$, $M(E_{1\alpha})=0.50$;
- для $0.0269 < \alpha < 0.0678$: $\Delta\alpha=0.0410$; $E_{1\alpha}=\{0.1; 0.2; \dots; 0.9; 1\}$, $M(E_{1\alpha})=0.55$;
- для $0.0678 < \alpha < 0.9731$: $\Delta\alpha=0.9053$; $E_{1\alpha}=\{1\}$, $M(E_{1\alpha})=1.00$.

Тогда, согласно (10) точечной оценкой нечёткого вывода E_1 будет число

$$F(E_1) = \frac{1}{0.9731} \int_0^{0.9731} M(E_{1\alpha}) d\alpha = \frac{1}{0.9731} (0.5 \cdot 0.0269 + 0.55 \cdot 0.0410 + 1 \cdot 0.9053) = 0.9673.$$

Аналогичными действиями были установлены точечные оценки нечётких выводов относительно уровней конфиденциальности и для остальных документов: $a_2 - F(E_2)=0.8534$; $a_3 - F(E_3)=0.7250$; $a_4 - F(E_4)=0.4823$; $a_5 - F(E_5)=0.3753$; $a_6 - F(E_6)=0.9655$; $a_7 - F(E_7)=0.8578$; $a_8 - F(E_8)=0.2981$; $a_9 - F(E_9)=0.5823$; $a_{10} - F(E_{10})=0.3756$. В результате, путём простого умножения этих значений на 100 получим соотношение итоговых оценок уровней конфиденциальности документов в масштабе отрезка $[0; 100]$.

Заключение

Традиционные методы установления степени конфиденциальности сведений, подлежащих включению в будущий документ, основаны на применении эвристических знаний лиц, ответственных за присвоение грифа конфиденциальности (в частности, системы предпочтений РКД). Однако процедура идентификации контекстных сведений и сама система предпочтений руководителя соответствующей службы очень сложны и, поэтому, требует перманентного сбора, хранения и оперативной обработки соответствующей информации, включая слабо структурированные данные, отражающие мнения экспертов относительно уровня конфиденциальности сведений, подлежащих включению в будущие документы. По средствам применения механизма нечёткого вывода, стало возможным, во-первых, классифицировать оценки влияния ФК сведений на уровень секретности документа, во-вторых, обобщать мнения экспертов о существующих ФК и, наконец, в-третьих, консолидировать выводы всех экспертов относительно уровня конфиденциальности сведений, подлежащих включению в будущий документ.

Сравнение полученных обоими методами результатов оценки уровней конфиденциальности рассмотренных альтернатив (документов) a_k ($k=1 \div 10$) представлен в табл.3, откуда видно, что итоговые оценки уровней конфиденциальности не везде совпадают, особенно, при сравнении уровней конфиденциальности по их наименованиям.

Таблица 3. Сравнительный анализ результатов оценки уровней конфиденциальности

Док-т (п/п)	Метод взвешенных оценок ФК			Метод нечёткого вывода		
	Итоговая оценка	Уровень конфиденциальности	Порядок	Итоговая оценка	Уровень конфиденциальности	Порядок
a_1	93.27	Строго конфиденциально	1	96.73	Строго конфиденциально	1
a_2	85.19	Конфиденциально: под особым контролем	3	85.34	Строго конфиденциально	4
a_3	77.22	Конфиденциально	5	72.50	Строго конфиденциально	5
a_4	58.27	Для служебного пользования	7	48.23	Конфиденциально: под особым контролем	7
a_5	47.90	Не нуждается в грифе конфиденциальности	9	37.53	Для служебного пользования	9
a_6	91.62	Строго конфиденциально	2	96.55	Строго конфиденциально	2
a_7	81.91	Конфиденциально: под особым контролем	4	85.78	Строго конфиденциально	3
a_8	37.78	Не нуждается в грифе конфиденциальности	10	29.81	Не нуждается в грифе конфиденциальности	10
a_9	68.09	Конфиденциально	6	58.23	Строго конфиденциально	6
a_{10}	50.84	Для служебного пользования	8	37.56	Для служебного пользования	8

Последнее объясняется разными подходами к формированию шкалы градации итоговых оценок уровней конфиденциальности. Тем не менее, ранжирование документов по итоговым оценкам, полученным на основе нечёткого подхода, вызывает большее доверие,

т.к. в данном случае при её формировании прослеживаются причинно-следственные связи между ФК и, собственно, уровнями конфиденциальности. Что касается порядков соотношений итоговых оценок, то они достаточно схожи, за исключением результатов, полученных для альтернатив a_2 и a_7 .

Литература

1. Алексенцев А.И. О классификации конфиденциальной информации по видам тайны. // Безопасность информационных технологий. – 1999, № 3. – С.65-71.
2. Алексенцев А.И. О составе защищаемой информации // Безопасность информационных технологий. – 1999, №2. – С.5-7.
3. Фатьянов А.А. Проблемы защиты конфиденциальной информации, не составляющей государственную тайну. // Информационное общество. – 1997, №1. – С.48-56.
4. Кострова В.Н. Оптимизация управления вузом на основе экспертно-мониторингового анализа структурно-функциональных компонентов образовательного процесса. / Автореф. дис. на соис. уч. степ. д-ра тех. наук / ВГТУ. – Воронеж, 2004. – 33 с.
5. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. // Математика. Новое в зарубежной науке: Пер. с англ., под ред. Н.Н. Моисеева и С.А. Орловского – М.: Мир, 1976. – 166 с.
6. Рзаев Р.Р. Аналитическая поддержка принятия решений в организационных системах. – Saarbruchen (Germany): Palmerium Academic Publishing, 2016. – 306 с.

Xülasə

Süleymanova A.N.

Gələcək sənədə daxil ediləcək məlumatların məxfilik dərəcəsinin qeyri-səlis təhlil əsasında qiymətləndirilməsi

Gələcək sənədə daxil ediləsi məlumatların məxfilik qrifinin adekvat səviyyəsinin vaxtında təyin edilməsi üçün zəruri olan zəminlər tədqiq edilir. Məxfilik faktorların qiymətləndirilməsi üçün uzaq mənbələrdən daxil olan müvafiq ekspert rəyləri tətbiq edilir və ilkin emaldan sonra qeyri-səlis çıxarış sistemi vasitəsilə onlar təsnif edilir. Sənədə məxfilik qrifinin verilməsi ekspert qiymətlərinin aqreqasiyası əsasında yerinə yetirilir.

Açar sözlər: məxfilik faktoru, ekspert rəyi, qeyri-səlis çoxluq, qeyri-səlis çıxarış sistemi.

Summary

Suleymanova A.N.

Assessment the confidentiality degree of information to be included in a future document on the base of fuzzy analysis

The prerequisites required as a prior condition for the timely establishment of an adequate level of the confidentiality of classified information to be included in future documents are considered. For assigning the confidentiality stamp, it is suggested to use expert conclusions from remote sources and for relevant data processing relative to classification of expert estimations and its following aggregation it is proposed using the fuzzy inference.

Key words: confidentiality factor, expert conclusion, fuzzy set, fuzzy inference system.