УДК 004.681 **Кравченко В.І., Левченко Є.Г.**

ВИКОРИСТАННЯ ТЕОРІЇ НЕЧІТКИХ МНОЖИН ДЛЯ ВИЗНАЧЕННЯ ВИТРАТ НА ЗАХИСТ ІНФОРМАЦІЇ

Розглядається система захисту інформації, яка містить два об'єкти з різною вразливістю. Відповідно до теорії нечітких множин утворені дві множини, одна з котрих ототожнюється з поставленою метою (виділена кількість ресурсів), а друга – з введеним обмеженням (допустимий рівень витоку інформації). Сформовані функції належності до утворених множин, які дають можливість розрахувати інтервали допустимих витрат на захист, котрі задовольняють поставленим умовам для кожного з об'єктів. Проведені розрахунки для різних функцій, визначені інтервали витрат для всієї системи.

Вступ. Протистояння двох сторін в інформаційній сфері ведеться в умовах невизначеності, коли дії суперника не можна передбачити, або в умовах ризику, коли їх можна оцінити з певною імовірністю. При цьому параметри і залежності, які формують математичну модель, знаходять на основі інтуїтивних уявлень за допомогою статистичних даних, а при їх відсутності – в результаті експертних оцінок. Величини, які становлять предмет пошуку і формують цільову функцію, визначаються наближено. При цьому здається доцільним звернутись до теорії нечіткої логіки і нечітких множин, які надають додатковий інструмент пошуку рішення в розпливчатих умовах [1-2].

Постановка задачі. Поставимо мету: визначити допустимий інтервал значень ресурсів захисту, в якому витік інформації не буде перевищувати заданий рівень. Для прикладу розглянемо інформаційну систему, яка складається з двох об'єктів з різною вразливістю. Використаємо методику Белмана-Заде [1], в якій процес прийняття рішень ведеться в умовах невизначеності. При цьому і шукана величина, і обмеження, які формують цільову функцію, задаються нечіткими множинами.

Результати дослідження. Сформулюємо задачу наступним чином:

а) нечітка мета: «ресурс захисту *y* має бути близьким до y_0»;

б)нечітке обмеження: «частка втраченої інформації *f* не повинна значно перевищувати f_0».

Величини, які входять в ці умови, нормовані до кількості інформації.

Відповідно до теорії нечітких множин введемо такі поняття:

$\mathrm{Y} = \{y\}$ - множина альтернатив;

$G(y)$ - нечітка множина в Y, яка ототожнюється з поставленою метою;

$C(y)$ - нечітка множина в Y, яка ототожнюється з введеним обмеженням.

Функції належності до введених нечітких множин сформуємо у вигляді гладких аналітичних функцій. Нечітке число, яке входить в формулювання поставленої мети, представимо нечіткою множиною $G(y)$ з такою функцією належності:

$$\mu_G(y) = \frac{a}{a + b * (y - y_0)^2}. \qquad (1)$$

Значення параметрів в (1) відображають рівень строгості поставленої умови відносно *y* і проявляються у формі залежності $\mu_G(y)$. Зокрема, чим більше *b*, тим вужча крива $\mu_G(y)$ і тим більш строго повинна виконуватись нечітка мета.

Вид цієї функції належності до нечіткої множини *C* визначається залежністю *f(x,y)* частки вилученої інформації від співвідношення ресурсів нападу і захисту. Ці залежності можна виражати з допомогою дробно-лінійних або дробно-нелінійних функцій [3]. На першому етапі будемо використовувати дробно-лінійні функції:

$$f(x, y) = \frac{x/y}{x/y + c} = \frac{1}{1 + c(y/x)}$$

Задавши значення *x,* переходимо від функції *f(x,y)* до *f(y):*

$$f(y) = \frac{1}{1 + \tilde{c}y}\ ,\ \text{де}\ \tilde{c} = \frac{c}{x}$$

Функцію належності до нечіткої множини $C(y)$ сформуємо у вигляді:

$$\mu_C(y) = \frac{cy}{1 + cy} \tag{2}$$

Монотонне зростання цієї функції свідчить про те, що зі збільшенням витрат втрати інформації зменшуються, введене нечітке обмеження виконується з більшою певністю, що й відображається зростанням функції належності.

Зауважимо, що залежності *f(y)* і $\mu_C(y)$ мають протилежний характер: при зростанні *y* функція *f(y)* спадає (при $y \to \infty$ до 0), що відображається зростанням $\mu_C(y)$ (при $y \to \infty$ до 1).

Параметри *a, b, c*, які входять в (1), (2), визначаємо з наступних міркувань. Параметр *a* в (1) не має суттєвого впливу на функцію $\mu_G(y)$. Цей параметр взагалі може бути об'єднаний з *b* і вводиться для зручності – щоб позбутись занадто великих значень *b.* Параметр *b,* який впливає на ширину лінії $\mu_G(y)$, визначається рівнем толерантності менеджменту до поставленої мети. Параметр *c* в (2) також визначається з суб'єктивних міркувань, а саме – заданим рівнем строгості виконання нечіткого обмеження.

Нашою метою є встановлення інтервалу допустимих значень виділених на кожний об'єкт ресурсів захисту. Інтервал визначається по заданому рівню певного результуючого показника, який враховує ступінь виконання обох нечітких умов. Цей ступінь задається менеджментом і залежить від загальної кількості ресурсів і допустимого рівня ризику. В нашому розгляді результуючим показником є $\mu(y) = \sqrt{\mu_G(y) \cdot \mu_C(y)}$, а фактори, які впливають на його величину (їх вплив і підлягає дослідженню) є:

1) форма функції $\mu_G(y)$, зокрема її ширина (параметр *b* в (1)) і ступінь асиметрії (вона проявляється далі у більш складних залежностях $\mu_G(y)$);
2) положення функції $\mu_G(y)$ на осі *y* (параметр y_0 в (1));
3) кривизна функції $\mu_C(y)$ (параметр *c* в (2)).

Вплив цих факторів видно з рис. 1-4. У всіх варіантах розрахунків задані такі значення y_0: для першого об'єкта $y_0^{(1)} = 0{,}08$, для другого - $y_0^{(2)} = 0{,}11$ (8% і, відповідно, 11% від вартості інформації на об'єкті, котру для двох об'єктів вважаємо однаковою).

Надалі використані наступні функції належності.

Варіант 1.

Для першого об'єкта:

$$\mu_G^{(1)}(y) = \frac{0.08}{0.08 + 100(y - 0.08)^2} \qquad \mu_C^{(1)}(y) = \frac{10y}{1 + 10y} \tag{3}$$

Для другого об'єкта:

$$\mu_G^{(2)}(y) = \frac{0.11}{0.11 + 100(y - 0.11)^2} \qquad \mu_C^{(2)}(y) = \frac{16y}{1 + 16y} \tag{4}$$

Параметр b=100 в $\mu_G(y)$ вибраний довільно і надалі буде змінюватись. Параметри c=10 в $\mu_C^{(1)}(y)$ і c=16 в $\mu_C^{(2)}(y)$ обумовлені використанням залежностей:

$$f_1(x,y)=\frac{x/y}{x/y+10} \qquad f_2(x,y)=\frac{x/y}{x/y+16} \tag{5}$$

Величини c_1=10, c_2=16 в цих виразах вибрані такими, що при $x/y=1$ дають значення f_1 = 0,091, f_2 = 0,058, які, на нашу думку, можуть відображати реальні ситуації. При заданому значенні f_0 = 0,1 (допустима кількість втраченої інформації – 10%) з (5) маємо: для першого об'єкта $x/y=1{,}1$ і при обраному значенні x=0,2 (ресурс нападу складає 20% від вартості інформації на об'єкті), одержуємо y = 0,18 і з (3) $\mu_C{}^{(1)}=0{,}64$, для другого об'єкта $x/y=1{,}78$; y = 0,112 і з (4) $\mu_C{}^{(2)}=0{,}66$.

Інтервал допустимих значень $\mu^{(k)}$ (k – номер об'єкта) знаходимо з умови $\sqrt{\mu_G{}^{(1)}(y)\cdot\mu_C{}^{(1)}(y)}=\sqrt{\mu_G{}^{(2)}(y)\cdot\mu_C{}^{(2)}(y)}=\mu$. Значення μ в подальших розрахунках приймемо рівним $\mu=0{,}33$.

Результати розрахунків по варіанту 1 приведені на рис.1.

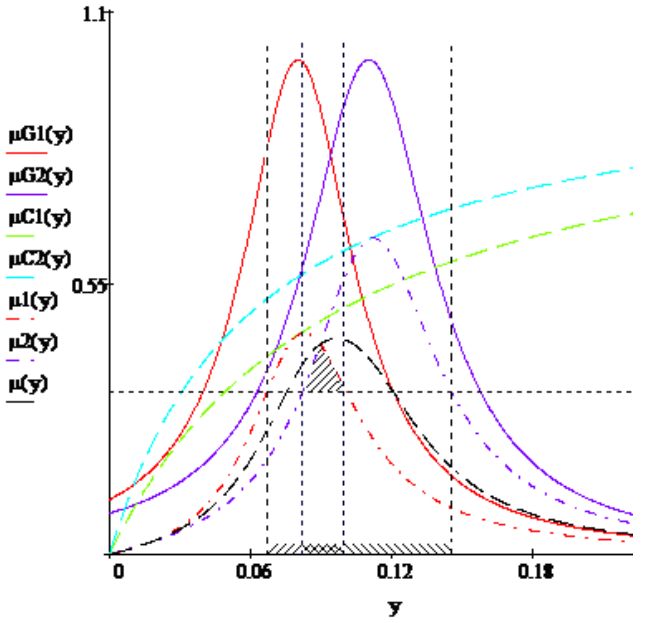

Рис. 1. Функції належності при симетричному характері $\mu_G(y)$

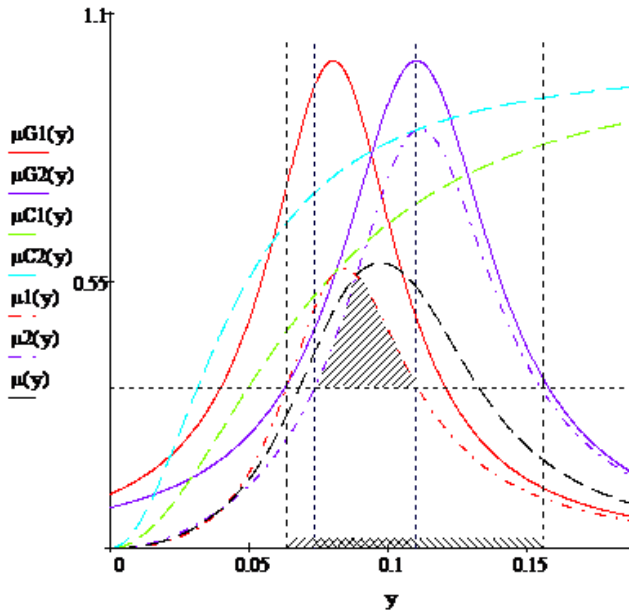

Рис. 2. Функції належності при дробно-нелінійному характері $\mu_C(y)$

Варіант 2.

$$\mu_G{}^{(1)}(y)=\frac{0.08}{0.08+400(y-0.08)^2} \qquad \mu_C{}^{(1)}(y)=\frac{10y}{1+10y}$$

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+400(y-0.11)^2} \qquad \mu_C{}^{(2)}(y)=\frac{16y}{1+16y}$$

Варіант 3.

$$\mu_G^{(1)}(y)=\frac{0.08}{0.08+100(y-0.08)^2} \qquad \mu_C^{(1)}(y)=\frac{32y}{1+32y}$$

$$\mu_G^{(2)}(y)=\frac{0.11}{0.11+100(y-0.11)^2} \qquad \mu_C^{(2)}(y)=\frac{64y}{1+64y}$$

Варіант 4.

$$\mu_G^{(1)}(y)=\frac{0.08}{0.08+400(y-0.08)^2} \qquad \mu_C^{(1)}(y)=\frac{32y}{1+32y}$$

$$\mu_G^{(2)}(y)=\frac{0.11}{0.11+400(y-0.11)^2} \qquad \mu_C^{(2)}(y)=\frac{64y}{1+64y}$$

Варіант 5.

$$\mu_G^{(1)}(y)=\frac{0.08}{0.08+100(y-0.08)^2} \qquad \mu_C^{(1)}(y)=\frac{200y^2}{1+200y^2}$$

$$\mu_G^{(2)}(y)=\frac{0.11}{0.11+100(y-0.11)^2} \qquad \mu_C^{(2)}(y)=\frac{500y^2}{1+500y^2}$$

На рис.2 приведені для порівняння результати розрахунків по варіанту 5. Результати цих і подальших розрахунків зведені в табл.1.

Таблиця 1

Допустимі інтервали ресурсів захисту

№ варіанта	Об'єкт №1				Об'єкт №2				Інтервал перекриття
	y_1	y_2	$\Delta y^{(1)}$	$\Delta y^{(1)} / y_0^{(1)}$	y_1	y_2	$\Delta y^{(2)}$	$\Delta y^{(2)} / y_0^{(2)}$	Δy
1	0.067	0.099	0.032	0.400	0.081	0.144	0.063	0.572	0.018
2	0.072	0.089	0.017	0.212	0.094	0.126	0.032	0.290	0.005
3	0.053	0.112	0.059	0.737	0.069	0.153	0.084	0.763	0.043
4	0.065	0.095	0.032	0.400	0.089	0.131	0.042	0.381	0.006
5	0.063	0.109	0.046	0.575	0.073	0.154	0.081	0.736	0.036
6	0.065	0.122	0.057	0.712	0.081	0.172	0.091	0.827	0.041
7	0.021	0.122	0.101	1.260	0.011	0.091	0.080	0.720	0.070

Розглянемо тепер випадок, коли функції належності $\mu_G(y)$ мають асиметричний характер, причому більш м'які вимоги ставляться до значень $y < y_0$ (рис.3). Це відповідає ситуації, коли менеджмент не відчуває суворих обмежень в ресурсах і основні вимоги ставить до зменшення ризику втрати інформації.

Варіант 6.

$$\mu_G^{(1)}(y)=\frac{y+10y^2}{y+10y^2+100(y-0.08)^2} \qquad \mu_C^{(1)}(y)=\frac{10y}{1+10y}$$

$$\mu_G^{(2)}(y)=\frac{y+5y^2}{y+5y^2+100(y-0.11)^2} \qquad \mu_C^{(2)}(y)=\frac{16y}{1+16y}$$

Наступний варіант відноситься до ситуації, коли менеджмент має протилежні пріоритети: втрачати значні кошти не має сенсу, оскільки ми знаходимось поблизу зони, в якій інвестиції в захист інформації недоцільні. В цьому випадку змінюється формулювання нечіткої мети: «ресурс захисту y повинен бути якомога меншим». Функції належності до множини $G(y)$ задаємо у вигляді спадаючих експоненціальних функцій (рис.4).

Варіант 7.

$$\mu_G{}^{(1)}(y) = e^{-7y} \qquad \mu_C{}^{(1)}(y) = \frac{30y}{1+30y}$$

$$\mu_G{}^{(1)}(y) = e^{-10y} \qquad \mu_C{}^{(1)}(y) = \frac{50y}{1+50y}$$

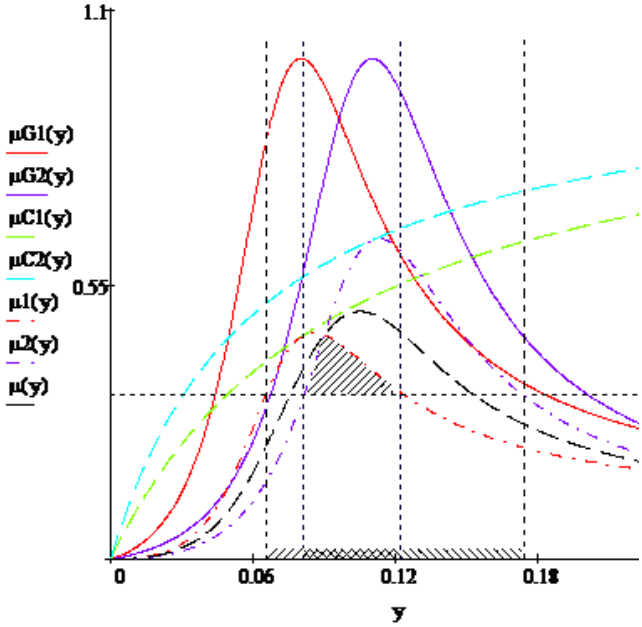

Рис. 3. Функції належності при асиметричному характері $\mu_G(y)$

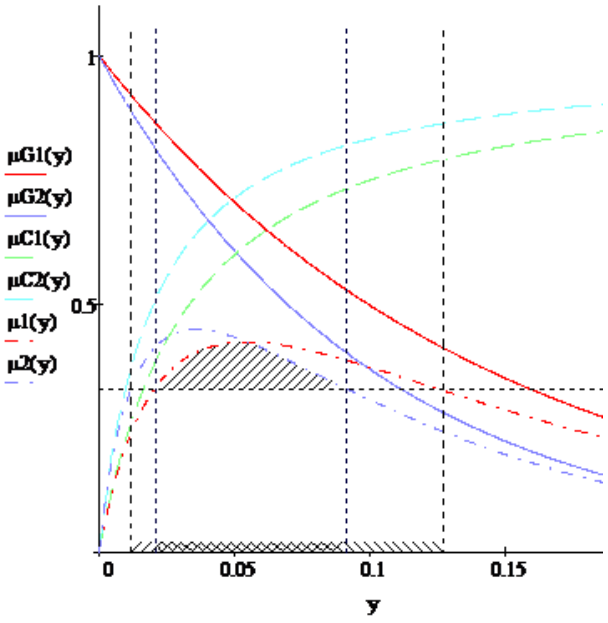

Рис. 4. Функції належності при експоненціальному характері $\mu_G(y)$

Приведені результати дають можливість оцінити наступні показники:

1) інтервали $\Delta y^{(1)}$, $\Delta y^{(2)}$ допустимих значень, які відповідають поставленим умовам, для першого і другого об'єктів (вони позначені лівонаправленою для першого об'єкта і правонаправленою для другого штриховкою);
2) відносні інтервали допустимих значень $\Delta y^{(1)} / y_0^{(1)}$, $\Delta y^{(2)} / y_0^{(2)}$;
3) інтервали перекриття Δy, на якому співпадають значення y для першого і другого об'єктів (і перекриваються штриховки на рисунках).

Якщо інтервали $\Delta y^{(k)}$ визначають допуски у виділенні ресурсів захисту на кожний з об'єктів, то інтервал перекриття має значення у випадку, коли різні об'єкти виділяють однакові засоби захисту з однаковою вартістю.

Висновки, які можна зробити з приведених результатів, підтверджують логічні передбачення і можуть служити їх кількісною ілюстрацією:

1) при посиленні вимог до виконання поставлених умов (звуження кривої $\mu_G(y)$, - перехід від варіанта 1 до варіанта 2, - і зниженні значень $\mu_C(y)$) інтервали $\Delta y^{(k)}$ звужуються – з $\Delta y^{(1)} = 0.032$, $\Delta y^{(2)} = 0.063$ у варіанті 1 до $\Delta y^{(1)} = 0.017$, $\Delta y^{(2)} = 0.032$ у варіанті 2;
2) перехід від дробно-лінійної до дробно-нелінійної функції належності $\mu_C(y)$ (варіант 5) приводить до розширення інтервалів $\Delta y^{(k)}$, а саме - до $\Delta y^{(1)} = 0.046$, $\Delta y^{(2)} = 0.081$;
3) при послабленні вимог до можливої кількості виділених ресурсів (варіант 6) інтервали $\Delta y^{(k)}$ пересуваються в область більших значень y, а при їх посиленні (варіант 7) – в область менших значень: $\Delta y^{(1)} = 0.057$, $\Delta y^{(2)} = 0.091$ у варіанті 6, $\Delta y^{(1)} = 0.101$, $\Delta y^{(2)} = 0.080$ у варіанті 7.

Висновки. Підводячи підсумки, зазначимо наступне. При моделюванні протистояння в сфері інформаційної безпеки, оперуючи неформальними (нечіткими) поняттями, ми прагнемо описати ці поняття деякими функціями розподілу, подібними імовірнісним функціям і далі використовуємо їх як точні, не дивлячись на їх «нечітку» природу. Наявність засобів теорії нечітких множин дозволяє побудувати математичну модель.

Приведена методика може бути застосована для розрахунку допустимих витрат на захист інформації в об'єктах, які відрізняються кількістю інформації, вразливістю та вимогами до допустимого рівня витрат, а також показниками, які використовуються при формуванні нечітких множин.

Література

1. Zadeh L.A. Fuzzy sets. Information and control. – 1965, Vol. 8, p. 338-353.
2. Bellman R.E., Zadeh L.A. Decision making in a fuzzy environment.-Managing Science., 1970. V.17, 4, p.141-164.
3. Левченко Є.Г. Оптимізаційні задачі менеджменту інформаційної безпеки // Рабчун А.О./ НТЖ «Сучасний захист інформації». – 2010.-№1.-с.16-23.

Надійшла: 11.03.11

Рецензент: д.т.н., проф. Петров О.С.