УДК 004.681

МАТЕМАТИЧНІ МОДЕЛІ ЕКОНОМІЧНОГО МЕНЕДЖМЕНТУ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Є.Г. ЛЕВЧЕНКО, М.В. ДЕМЧИШИН, А.О. РАБЧУН

Проведено порівняльний аналіз кількісних моделей економічного менеджменту інформаційної безпеки. Встановлено умови, за яких цільові функції розглянених моделей співпадають або досить близькі. Виконані розрахунки ілюструють методику визначення кількості втраченої інформації в залежності від співвідношення ресурсів нападу і захисту.

ВСТУП

Основним завданням економічного менеджменту інформаційної безпеки є пошук оптимальних рішень. При цьому ми стикаємось з низкою проблем, які мають як об'єктивний, так і суб'єктивний характер. Головна з них — побудова математичної моделі, яка включає:

- вибір критеріїв оптимальності, які враховують вибрані пріоритети, тобто важливість для підприємства таких показників, як ризик втрати інформації, сума витрат на її захист, рентабельність витрат тощо;
- визначення параметрів розрахунку та функціональних залежностей, які входять в математичну модель.

Основні задачі, що стоять перед менеджментом:

- визначення оптимальної кількості ресурсів захисту, які мінімізують сумарні витрати, що включають в себе потенційні втрати від витоку інформації і витрати на її захист із врахуванням відповідних вагових коефіцієнтів;
- оптимізація розподілу ресурсів між об'єктами, які містять різні обсяги інформації, характеризуються різним рівнем вразливості та певним ступенем корельованості, а також між окремими ступенями захисту;
- визначення оптимального розподілу ресурсів в умовах комплексного протистояння в конкурентній боротьбі, коли кожна сторона захищає свою інформацію й одночасно спрямовує свої зусилля на здобуття інформації конкурента, причому частина ресурсів може бути спрямована на розвідку;
- визначення зміни станів інформаційної безпеки з часом із врахуванням можливих дій суперників;
- розробка методики управління ресурсами в динамічному режимі, в якій враховано наведені ситуації та показники.

Перший крок у дослідженні процесів протистояння — це розробка аналітичної математичної моделі, яка має задовольняти двом суперечливим вимогам: максимально відображати найважливіші аспекти протистояння двох сторін в інформаційній сфері й одночасно уникати зайвого ускладнення, яке могло б ускладнити отримання практичних результатів.

Побудову моделі можна поділити на декілька етапів.

- Визначення параметрів та показників системи захисту інформації: кількість l об'єктів захисту; обсяг g_k інформації на кожному об'єкті (k — номер об'єкта); початкова вразливість υ_k об'єкта; виділений ресурс Y захисту; відношення до ризику.
- Оцінка дій суперника: характер атак (націлені, ненацілені); виділений ресурс X нападу; імовірність p_k нападу на об'єкт; імовірності виділення нападом ресурсів x_k на кожний об'єкт.
- Формування цільової функції, яке включає вибір цільового показника і незалежної змінної та встановлення між ними функціональної залежності: цільовий показник (кількість втраченої інформації; сумарні витрати ресурсів, які включають втрати від витоку інформації та витрати на її захист; ефективність інвестування, яку визначаємо як частку двох величин — зменшення обсягу втраченої інформації та витрат на захист); незалежні змінні (ресурси нападу і захисту — x_k та y_k; динамічна вразливість $\upsilon_k(t)$); вид функціональної залежності цільового показника від незалежної змінної (степенева, показникова).

Мета роботи — порівняльний аналіз моделей економічного менеджменту інформаційної безпеки.

РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

Під час побудови математичної моделі ключовим питанням є формування цільової функції. Перша відома нам кількісна модель наведена в [1], де подана проблема була піддана ґрунтовному аналізу. У моделі Гордона-Лоеба (ГЛ) цільова функція визначає зменшення втрат в інформаційній системі (завдяки внесенню інвестицій) з відрахуванням витрат y на її захист (в [1] ці витрати позначені через z). У функцію як параметр входить вразливість υ об'єкта, тобто імовірність того, що напад буде успішним при $y = 0$. Авторами [1] запропоновано два широких класи функцій $S(y;\upsilon)$, які визначають імовірність втрати інформації:

$$S^I(y;\upsilon) = \frac{\upsilon}{(\alpha y + 1)^{\beta}}; \tag{1}$$

$$S^{II}(y;\upsilon) = \upsilon^{\alpha y + 1}, \tag{2}$$

де параметри $\alpha > 0$, $\beta \geq 1$ характеризують продуктивність інформаційної безпеки. Цільова функція має вигляд:

$$E(y) = [\upsilon - S(y,\upsilon)]L - y, \tag{3}$$

де L — потенційні втрати інформації при здійсненні нападу.

У (3) перша складова (υL) визначає кошти, втрачені в результаті нападу за відсутності системи захисту інформації (СЗІ), друга $(S(y;\upsilon)L)$ — при введенні СЗІ, третя (y) — інвестиції в СЗІ. Загалом (3) визначає кошти, збережені завдяки введенню СЗІ (в економічній термінології — прибуток від інвестицій).

Метою аналізу в [1] є визначення оптимальних витрат y^0 при різних значеннях вразливості. Показником оптимальності є максимум прибутку від

інвестицій, що виражається умовою $E'_y(y) = 0$. Показано, що вид залежності $y^0(\upsilon)$ відрізняється для двох класів функцій $S(y;\upsilon)$, і для розробки рекомендацій із визначення раціональної кількості інвестицій, крім вибору виду функції $S(y;\upsilon)$, необхідно встановити рівень вразливості об'єкта.

Останнім часом з'явилась низка робіт [2–5], спрямованих на розвиток моделі ГЛ. Зокрема, в [2] зосереджена увага на тому, що інвестиції в інформаційну безпеку можуть не тільки зменшувати можливі втрати, але й відлякувати потенційного порушника і в результаті зменшувати імовірність загрози. Розглядаючи ці явища (зменшення втрат і зменшення загрози), можна виокремити три варіанти їх взаємодії: відсутність впливу, позитивний вплив і негативний вплив. Для детальнішого дослідження цього питання в [2] введені поняття продуктивності зменшення вразливості й продуктивності зменшення загрози, а також простору продуктивності, який об'єднує ці показники. В залежності від їх значень випливають висновки щодо вибору стратегій, які забезпечують оптимальні витрати на захист інформації.

У [3] використані дещо змінені залежності $S(y;\upsilon)$:

$$S^I(y;\upsilon) = \frac{p\upsilon}{\alpha y + 1}, \tag{4}$$

$$S^{II}(y;\upsilon) = p \cdot \upsilon^{\alpha y + 1}, \tag{5}$$

де p — імовірність здійснення нападу. Предметом дослідження в [3] є визначення оптимального розподілу ресурсів нападу між об'єктами, на які здійснюють численні напади, в умовах обмеження бюджету й у випадку, коли такі обмеження відсутні. Показано, що існують граничні значення вразливості, за межами яких інвестиції недоцільні, значно простіше відшкодовувати втрати. Численні припущення під час побудови моделі [3], за думкою авторів, обмежують застосування цієї моделі.

В [4] проведено деяке узагальнення сімейства функцій $S(y;\upsilon)$ і запропоновано їх нові, складніші формулювання. Проте ступінь їх корисності ще варто встановити. В [5] зроблено спробу здійснення динамічного аналізу інвестування в інформаційну безпеку, причому порушення інформації розглядається як випадковий процес, який характеризується показниками дрейфу та волатильності. Наведено результати розрахунків на основі моделі ГЛ, в яких використано гіпотетичні значення параметрів.

У [6] запропоновано інший підхід до цієї проблеми. Математична модель [6] передбачає використання цільової функції $i(x;y)$, де i — віднесена до загальної кількості інформації вартість втраченої інформації, x та y — ресурси нападу i, відповідно, захисту. Ця функція в загальних ознаках має вигляд:

$$i(x,y) = \sum_{k=1}^{l} i_k(x,y) = \sum_{k=1}^{l} g_k p_k q_k(x,y) f_k(x,y), \tag{6}$$

де $k = \overline{1,l}$ — номер об'єкта; g_k — обсяг інформації на об'єкті; p_k — імовірність нападу на об'єкт; $q_k(x;y)$ — щільність імовірності виділення напа-

дом ресурсів x на k-й об'єкт; $f_k(x;y)$ — залежність частки втраченої інформації від співвідношення x та y, яку можна розглядати як імовірність втрати інформації при заданих значеннях x та y.

В якості залежностей $f(x,y)$ запропоновано два класи функцій:

$$\text{степеневі} \quad f(x,y)=\frac{a(x/y)^n}{b(x/y)^n+c} \tag{7}$$

$$\text{та показникові} \quad f(x,y)=d(1-e^{-m(x/y)^n}), \tag{8}$$

де параметри a,b,c,d,n,m приймають додатні значення і визначають положення та нахил кривих.

В [6] запропоновано два можливих види залежностей $q(x)$ у вигляді $q(x)=Nx^n e^{-h^2x^2}$: розподіл Максвела $q_M(x)=Nx^2e^{-h^2x^2}$ і розподіл Релея $q_P(x)=Nxe^{-h^2x^2}$, де N — нормовочний коефіцієнт, а константи n,h визначають положення максимуму залежності і ступінь її асиметрії. У співставленні цих розподілів суттєва для нас їх відмінність полягає в тому, що для $q_M(x)$ в початковій області $x \gtrsim 0$ опуклість направлена вниз, а для $q_P(x)$ — вгору.

Порівнюючи зазначені моделі, спробуємо окреслити їх відмінності та спільні ознаки.

- У моделі ГЛ за мету ставиться оптимізація витрат y^0 на захист інформації, у моделі [6] — оптимізація розподілу ресурсів y_k між окремими об'єктами.
- Цільова функція ГЛ визначає зменшення втрат від можливого витоку інформації з відрахуванням витрат на її захист, у моделі [6] — вартість втраченої інформації.
- У моделі ГЛ основним показником системи захисту інформації, який входить у цільову функцію та впливає на оптимальне значення y^0, є вразливість υ, у моделі [6] — співвідношення ресурсів нападу і захисту (x та y).
- Функції, які визначають імовірність втрати інформації, у моделі ГЛ задаються виразами (1), (2), у моделі [6] — виразами (7), (8).
- Імовірність нападу й імовірність втрати інформації в моделі ГЛ входять у цільову функцію у вигляді параметрів α та β, які є мірою продуктивності витрат, у моделі [6] імовірність нападу задається в явному вигляді, а імовірність втрати інформації при здійсненому нападі — залежностями (7), (8).
- У моделі ГЛ величини, які входять у розрахунок, даються в абсолютному вимірі, у моделі [6] — у відносному.
- Спільним в обох моделях є те, що параметри розрахунку (υ, α, β та, відповідно, p, q) і функції, які визначають залежність частки вилученої інформації від вкладених ресурсів ($S(y;\upsilon)$ та $f(x;y)$), не можуть бути встановлені точно, а знаходяться в результаті аналізу статистичних даних, а в разі їх відсутності — на основі експертних оцінок.

Слід зазначити, що оптимізація розподілу ресурсів між об'єктами певною мірою зумовлена можливими стратегіями нападу. Напади суперника можуть бути ненаціленими (шкідливе програмне забезпечення, віруси, фішинг, спам тощо) і націленими (хакерські атаки на банківську базу даних з метою вилучення коштів або вибір об'єкта нападу в системі, яка містить певну кількість об'єктів). Характер нападів може бути зумовлений, в одних випадках, — цільовою спрямованістю зловмисника, в інших, — наявною кількістю ресурсів. Останній варіант спостерігається, зокрема, коли об'єкти однотипні, і конкурент розглядає доцільність розподілу обмежених ресурсів між окремими об'єктами. Націлені атаки трапляються рідше, ніж ненаціленні, проте їх наслідки можуть бути серйознішими для підприємства, що, звичайно, слід враховувати під час розподілу ресурсів захисту. В [3] зроблено припущення, що клас функцій (4) краще описує націлені атаки, а клас (5) — ненаціленні.

Порівняємо тепер залежності $S(y;\upsilon)$ та $i(x;y)$, які виражають практично одну і ту ж величину. За цього порівняння величину x у виразі $i(x;y)$ вважатимемо сталою, і для спрощення всі величини, крім $f(x;y)$, покладемо рівними одиниці. Таким чином, ми зведемо задачу до порівняння залежностей $S(y;\upsilon)$ (де υ — параметр) та $f(x;y)$ (x — параметр). Поклавши $x=1$, функція $f(x;y)$ (7) матиме вигляд:

$$f(y)=\frac{a}{b+cy^n}. \qquad (9)$$

Порівнюючи її із залежністю $S^I(y;\upsilon)=\dfrac{\upsilon}{(\alpha y+1)^\beta}$, можемо зробити такі висновки. При $\alpha=\beta=1$ в (1) та $a=\upsilon$, $b=c=n=1$ у (9) ці залежності співпадають повністю: $f(y,\upsilon)=S^I(y;\upsilon)=\dfrac{\upsilon}{y+1}$. При $a=\upsilon$, $b=1$, $c=\alpha$, $n=\beta$ маємо замість (9): $f(y,\upsilon)=\dfrac{\upsilon}{\alpha y^\beta+1}$. Значення цієї величини перевищують значення (1) при всіх α, β та y, а при $\beta=1$ співпадають. Таким чином, модель [6] із цільовою функцією (7) тим ближча до моделі ГЛ із цільовою функцією (1), чим ближчі α та β у (7) до одиниці.

Зазначимо, що чим більше значення n у (7), тим більший інтервал Δy, в якому $f(y)$ залишається максимальним («поличка» в області $y\gtrsim 0$ на рис. 1, *а*), де $f(y)=0{,}9/(1+cy^n)$. Таким чином, величину n можна сприймати як один із показників вразливості системи. Криві $f(y)$ з різними значеннями n перетинаються в точці $y=1$. Положення цієї точки по осі $f(y)$ визначається величиною c. Величина n впливає на кривизну залежностей $f(y)$, причому їх опуклість в області $y=0..1$ при $n>1$ направлена вгору, при $n\le 1$ — вниз. Така різноманітність форми надає певну свободу дій після накопичення достатньої кількості статистичних даних, на основі яких з'явиться можливість встановити форми залежності $f(y)$ для реальних ситуацій.

У залежностях $S^I(y)$ «поличка» з $S = \text{const}$ відсутня, як і точка перетину кривих, але криві за формою досить схожі (рис. 1, *б*), де $S(y) = 0{,}9/(\alpha y + 1)^{\beta}$. Вплив параметрів α та β на їх крутизну видно з рис. 1

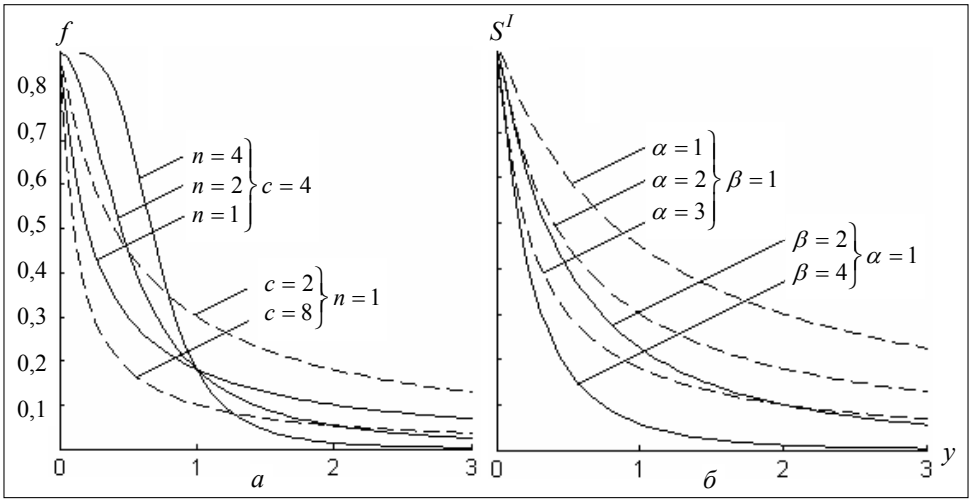

Рис. 1. Порівняння степеневих функцій $f(x;y)$ та $S^I(y;\upsilon)$

Порівняння залежностей $S^{II}(y;\upsilon)$ та $f(x;y)$ викликає ще більший інтерес, оскільки в [2], спираючись на статистичні дані [7], зазначається, що функції $S^{II}(y;\upsilon)$ знайшли своє емпіричне підтвердження. Хід кривих $f(x;y)$ під час використання показникових функцій (8) у вигляді $f(y) = 0{,}9(1 - e^{-m/y^n})$, де $x = 1$, а $\upsilon = 0{,}9$ (при $\upsilon = 1$ (2) втрачає сенс) видно з рис. 2, *а*. Для цих кривих також характерна «поличка» і точка перетину кривих із різним n при $y = 1$, в якій значення $f(y)$ визначається величиною m. Параметр n впливає на крутизну $f(y)$. Вплив параметра α на хід кривих $S^{II}(y;\upsilon)$ під час використання залежності (2) показано на рис. 2, *б*, де

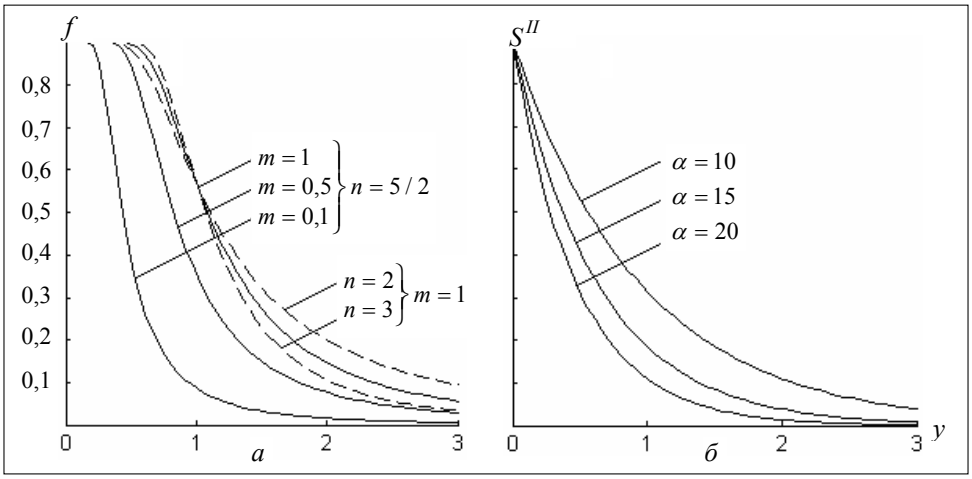

Рис. 2. Порівняння показникових функцій $f(x;y)$ та $S^{II}(y;\upsilon)$

$S(y,\upsilon)=\upsilon^{\alpha y+1}$, $\upsilon=0{,}9$. Для більш наочного співставлення результатів, розрахованих по залежностям (1) і (7) та, відповідно, (2) і (8) на рис. 3 наведене попарне порівняння цих залежностей, де суцільними і штриховими лініями зображено функції $f(y)$ при різних значеннях параметрів: на рис. 3, *а* $f(y)=0{,}9/(1+cy^n)$, на рис. 3, *б* $f(y)=0{,}9/(1-e^{-m/y^n})$, а штрих-пунктирними — функції $S(y)=0{,}9(\alpha y+1)^{\beta}$ на рис. 3, *а* та $S(y)=0{,}9^{\alpha y+1}$ на рис. 3, *б*.

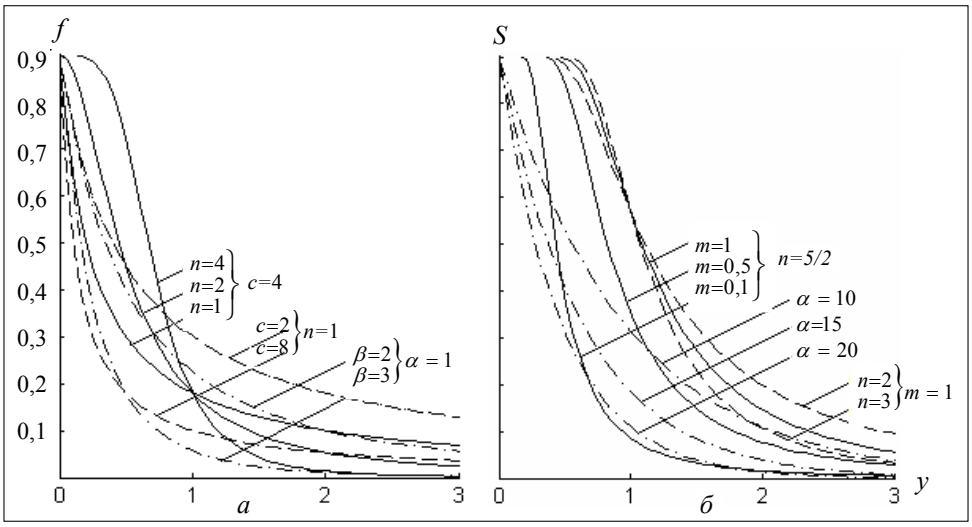

Рис. 3. Попарне порівняння степеневих та показникових функцій $f(x;y)$ та $S(y;\upsilon)$

Спробуємо підібрати параметри в (8), а також у (7) так, щоб ці залежності були в максимальній степені схожі на (2). Покладемо $\upsilon=0{,}9$, а параметр α в (2) виберемо з таких міркувань. Вважатимемо, що ресурси захисту y (відношення коштів, виділених на захист, до вартості інформації) знаходиться в межах $0..0{,}2$, а кількість вилученої інформації (тобто значення функцій $S(y;\upsilon)$ та $i(x;y)$), яка є небезпечною для підприємства, становить 15–20 % і катастрофічною — 20–30 %. Для вилучення такої кількості інформації для суперника буде доцільним виділити ресурс $(x/y)<3$. При цьому вважаємо, що при потрійному перевищенні ресурсів нападу $((x/y)\approx 3)$ суперник здатен вилучити значну кількість інформації. Значення $S^{II}(y;\upsilon)$, які відповідають цим міркуванням, наведені в табл. 1 і досягаються при $\alpha=30$:

$$S^{II}(y;v)=0{,}9^{30y+1}. \qquad (10)$$

Таблиця 1. Значення величин $S^{II}(y;\upsilon)$, $f_1(x;y)$, $f_2(x;y)$ при різних y та $x=1$

y	0	0,3	0,6	0,9	1,2	1,5	1,8	2,1	2,4	2,7	3
$S^{II}(y;\upsilon)$	0,9	0,348	0,135	0,052	0,020	0,007	0,003	0,001	0,0005	0,0002	0,00007
$f_1(x;y)$	0,9	0,500	0,120	0,045	0,022	0,012	0,008	0,005	0,004	0,003	0,0023
$f_2(x;y)$	0,9	0,453	0,136	0,054	0,027	0,016	0,010	0,007	0,005	0,003	0,002

Близькі до наведених значень $S^{II}(y;\upsilon)$ величини (які практично співпадають у найбільш важливому інтервалі $y = 0{,}5..1$, що відповідає $(x/y) = 1..2$) досягаються для показникової функції (8) у вигляді

$$f_1(y) = 0{,}9(1 - e^{-0{,}04(1/y)^{5/2}}) \tag{11}$$

і степеневої функції (7)

$$f_2(y) = 0{,}9\frac{(1/y)^{5/2}}{(1/y)^{5/2} + 20}, \tag{12}$$

значення цих функцій подано в табл. 1.

У вибраному інтервалі значень y степенева функція (12) ближча до функції $S^{II}(y;\upsilon)$ (10), ніж показникова функція (11). Це дозволяє припустити, що за певних значень параметрів степеневі функції (7) теж будуть відповідати емпіричним залежностям.

Врахуємо тепер імовірності $q(x,y)$, які оберемо у вигляді розподілу Максвела і використаємо в трьох формах:

$$q_1(x,y) = 2{,}26\left(\frac{x}{y}\right)^2 e^{-\left(\frac{x}{y}\right)^2}, \quad q_2(x,y) = 0{,}68\left(\frac{x}{y}\right)^2 e^{-0{,}44\left(\frac{x}{y}\right)^2},$$

$$q_3(x,y) = 0{,}36\left(\frac{x}{y}\right)^2 e^{-0{,}25\left(\frac{x}{y}\right)^2},$$

де максимум функції досягається при значенні $x_{1m} = 1$, $x_{2m} = 1{,}5$ та $x_{3m} = 2$ відповідно, а нормовочний коефіцієнт визначається з умови $\int_0^3 q_j(x)dx = 1$, $j = \overline{1,3}$. Під час розрахунку кількості вилученої інформації використаємо (7) у формі залежності від x (тобто у виразі (7) покладемо y рівним одиниці). Значення $i_j(x) = q_j(x)f_2(x)$, де $f_2(x) = 0{,}9\frac{x^{5/2}}{x^{5/2} + 20}$, наведені в табл. 2.

Таблиця 2. Кількість вилученої інформації з врахуванням імовірності $q(x)$ виділення нападом ресурсів x

$i_j(x)$ \ x	$x = 1$	$x = 1{,}5$	$x = 2$
$i_1(x)$	0,0356	0,0584	0,0328
$i_2(x)$	0,0187	0,0619	0,0928
$i_3(x)$	0,0120	0,0503	0,1051
$\bar{i}(x)$	0,0221	0,0569	0,0769

Середні значення $\bar{i}(x)$ показують кількість вилученої інформації при кожному значенні x за умови, що приведені залежності $q_j(x)$ рівноімовірні. Якщо в результаті експертної оцінки виявиться, що одна із залежностей

$q_j(x)$ має більш ваговий коефіцієнт c_j, ніж інші, то розподіл $i(x) = \sum_j c_j i_j(x)$ зміститься, очевидно, у бік цієї залежності.

ВИСНОВКИ

Наведені значення дозволяють менеджменту підприємства зробити висновок щодо достатності виділених коштів чи доцільності їх збільшення. Це залежить, звичайно, від допустимих величин $i(x, y)$, які, у свою чергу, визначаються з суб'єктивної оцінки топ-менеджера та його схильності до ризику. Остання ознака змушує дослідників розробляти різні моделі для випадків суворого обмеження ризиків і певного рівня їх допустимості [3]. При цьому модель ГЛ вважається побудованою на базі нейтрального відношення до ризику, за якого розглядається лише кінцевий результат (рівень необхідних інвестицій) без врахування рівня ризику.

Зазначимо, що подібна задача — мінімізація сумарних втрат (від витоку інформації та витрат на її захист, що рівноцінно підходу [3]) — розглядалась у [8], де проаналізовано загальні властивості цільової функції, і на основі аналізу зроблено висновки щодо максимально доцільного відсотку витрат на захист інформації. Порівняння методики в [8] із наведеними моделями потребує окремого дослідження.

ЛІТЕРАТУРА

1. *Gordon L.A., Loeb M.P.* The Economics of Information Security Investment, ACM // Transactions on Information and System Security. — 2002. — **5**, № 4. — P. 438–457.
2. *Matsuura K.* Productivity Space of Information Security in an Extension of the Gordon-Loeb's Investment Model // The Seventh Workshop on the Economics of Information Security (WEIS), 25–28 June, Hanover, USA. — 2008. — http://weis2008.econinfosec.org/papers/Matsuura.pdf .
3. *Huang C.D., Hu Q., Behara R.S.* Economics of Information Security Investment in the Case of Simultaneous Attacks // Proceeding of the Fifth Workshop on the Economics of Information Security, 26–28 June, Cambridge, England. — 2006. — http://weis2006.econinfosec.org/docs/15.pdf .
4. *Willemson J.* Extending the Gordon and Loeb Model for Information Security Investment, The Fifth International Conference on Availability, Reliability and Security ARES 2010 // Instityte of Electrical and Electrinics Engineers. — 2010. — http://research.cyber.ee/~jan/publ/aresGL.pdf .
5. *Tatsumi K., Goto M.* Optimal Timing of Information Security Investment: A Real Options Approach // The Eighth Workshop on the Economics of Information Security (WEIS 2009), UK, 24–25 June, University College London. — 2009. — P. 211–228. — http://weis09.infosecon.net/files/112/paper112.pdf .
6. *Левченко Є.Г., Рабчун А.О.* Оптимізаційні задачі менеджменту інформаційної безпеки // НТЖ «Сучасний захист інформації». — 2010. — № 1. — С. 16–23.
7. *Liu W., Tanaka H., Matsuura K.* Empirical-Analysis Methodology for Information-Security Investment and Its Application to a Reliable Survey of Japanese Firms // Information Processing of Japan Digital Courier. — 2007. — **3**. — P. 585–599.
8. *Задірака В.К., Олексюк О.С., Смоленюк Р.П., Штабалюк П.І.* Фінансування витрат на захист інформації в економічній діяльності // Університетські наук. зап. — 2006. — № 3–4 (19–20). — С. 479–490.

Надійшла 27.04.2010