МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Левченко Є.Г., Швець В.А., Демчишин М.В.

# ЕКОНОМІКА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Навчальний посібник

Київ
«НАУ – друк»
2012

УДК 004.056.5:658(075.8)
ББК з 973.20-018.я77 У497.8я7
Л 38

Рецензенти:

В.П. Мартинюк – доктор економічних наук, доцент, завідувач кафедри фінансово-економічної безпеки Тернопільського національного економічного університету.

Д.В. Чірков – кандидат технічних наук, доцент, заступник директора Навчально-наукового інституту захисту інформації Державного університету інформаційно-комунікаційних технологій.

В.М. Богданов – кандидат технічних наук, доцент ВІТІ НТУУ «КПІ»

Л38 Левченко Є.Г., Швець В. А., Демчишин М.В.
Економіка інформаційної безпеки: Навчальний посібник. – К.: НАУ, 2012. – 225 с.

ISBN

Приведені математичні моделі економічного менеджменту інформаційної безпеки та методи розрахунку її показників. Численні приклади ілюструють застосування розроблених методик до оптимізації показників складних систем захисту інформації.

Для студентів, аспірантів і викладачів навчальних закладів освіти галуззі знань 1701 "Інформаційна безпека".

ББК з 973.20-018.я77 У497.8я7
ISBN 

**ВСТУП**

# ВСТУП

Економіка – це наука про найбільш ефективне використання обмежених ресурсів. Інформація – це відомості, використання яких підвищує імовірність досягнення поставлених цілей. Економіка інформаційної безпеки – це розділ науки про найбільш ефективне використання ресурсів в системах захисту інформації. Інакше кажучи, економіка інформаційної безпеки направлена на застосування методів оптимізації в економічних задачах менеджменту інформаційної безпеки.

Необхідність пошуку оптимальних рішень в різних галузях людської діяльності привела до виникнення нового розділу математики – дослідження операцій, що базується на єдності типів математичних моделей і методів рішення різноманітних задач. Під дослідженням операцій розуміють застосування математичних кількісних методів для обґрунтування рішень у всіх областях цілеспрямованої людської діяльності.

При дослідженні операцій ключову роль відіграє математична модель – умовний образ деякої системи, який з допомогою математичних методів відображає властивості об’єктів, їх взаємозв’язків і економічних процесів, котрі виникають при їх взаємодії. При цьому важливо дотримуватись системного підходу, який в економічних задачах інформаційної безпеки проявляється в тому, що система «напад – захист» розглядається у взаємодії (в нашому випадку – у протидії) її складових з врахуванням їх параметрів і характеристик.

Слід зазначити, що нашим завданням є лише кількісне обґрунтування оптимальних рішень. Прийняття рішень відноситься до компетенції топ-менеджера, або, в нашій термінології, – особи, яка приймає рішення, і залежить від пріоритетів, котрі вона надає окремим показникам.

# І. АНАЛІЗ СТАТИСТИКИ НАПАДІВ В СФЕРІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

В нашій країні статистика нападів на інформаційні системи практично відсутня, тому суттєвий інтерес викликає аналіз зарубіжних даних. Приведемо деякі дані американських, японських і австралійських організацій (в основному, це звіти Інституту комп'ютерної безпеки США, Міністерства праці і економіки Японії, низки австралійських організацій) [27-31].

В нашому аналізі розглянуті такі аспекти інформаційної безпеки (ІБ):

1) загрози та інциденти в сфері ІБ;
2) збитки від основних видів загроз і порушень;
3) витрати на захист інформації;

Серед приведених даних (ми вважаємо, що вони відображають повну картину) нашу увагу привернули такі показники:

1) за профілем діяльності на першому місці стоять фінансові установи;
2) питання ІБ являється актуальним навіть для невеликих підприємств – з числом працівників менше 100;
3) за рівнем доходу найбільшу увагу питанням ІБ приділяють підприємства з високим доходом – більше 1 млрд. дол.

Статистику загроз та інцидентів у сфері ІБ ілюструють наступні результати.

Таблиця 1.1

Кількість інцидентів у відсотках до числа респондентів.

| Кількість інцидентів | 2006 | | | 2007 | 2008 |
|---|---|---|---|---|---|
| | США (616)* | Японія (216) | Австралія (389) | США (487) | США (517) |
| 1-5 | 48% | 42% | 69% | 46% | 47% |
| 6-10 | 15% | 7% | 18% | 14% | 14% |
| >10 | 9% | 4% | 5% | 22% | 13% |
| Не можуть відповісти | 28% | 9% | 8% | 18% | 26% |
| Не було інцидентів | - | 38% | - | - | - |

*загальна кількість респондентів.

Як видно з табл.1.1 (на прикладі США) процентне співвідношення кількості інцидентів за останні роки майже не змінилось.

Наступна таблиця дає змогу оцінити частість різних видів атак і, відповідно, ступінь їх небезпеки.

Серед електронних атак на першому місці стоять віруси. Більше половини компаній у приведених країнах зазнають вірусних атак на протязі року. Причиною цього є недосконалість засобів захисту, які намагаються розпізнати в елементах нападу риси відомих атак, в той час, як порушники розробляють нові віруси і вводять їх в місця, де можна обійти антивіруси за обмежений проміжок часу.

Таблиця 1.2

Види атак та зловживань у відсотках до числа респондентів.

| Вид інциденту | 2006 | | | 2007 | 2008 |
|---|---|---|---|---|---|
| | США (616) | Японія (984) | Австралія (327) | США (436) | США (433) |
| Вірусні атаки | 65 | 67 | 64 | 52 | 50 |
| Інсайдерські зловживання мережевим доступом | 42 | 18 | 62 | 59 | 44 |
| Крадіжки портативних носіїв | 47 | 23 | 58 | 50 | 42 |
| Неавторизований доступ | 32 | 5 | 8 | 25 | 29 |
| Відмова в обслуговуванні | 25 | 11 | 18 | 25 | 21 |
| Спам | - | - | - | - | 21 |
| Бот-мережі | - | - | - | - | 20 |
| Порушення/втрата даних користувачів | - | - | - | - | 17 |
| Зловживання бездротовою мережею | 14 | 3 | 10 | 17 | 14 |
| Проникнення в систему | 15 | 4 | 7 | 13 | 13 |
| Фінансове шахрайство | 9 | 12 | 6 | 12 | 12 |
| Порушення/втрати інформації<br>-з мобільних носіїв<br>-з інших джерел | 9 | 2 | 14 | 8 | 9<br>4<br>5 |
| Зловживання при роботі з веб-додатками | 6 | 3 | 8 | 9 | 11 |

Розглянемо тепер втрати, які несуть компанії в результаті електронних атак. В 2009 р. середня вартість втрат на одну

компанію в США становила 289 тис. дол. Найчастіше траплялись такі типи зловживань: віруси – їх зафіксували 49% респондентів, інсайдерські атаки, які супроводжувались крадіжкою даних з портативних носіїв – 44%. Проте в 2007р. зафіксовано викид фінансового шахрайства, який вивів його на перше місце (табл. 1.3). Найпоширенішими видами фінансового шахрайства є незаконне присвоєння активів, махінації з фінансовою звітністю, хабарництво і корупція (особливо це стосується фінансових і страхових компаній і фірм, які працюють в комунікаційних та технологічних галузях).

Таблиця1.3

Середні збитки від різних видів атак на одного респондента (в доларах США)

| Вид інциденту | 2006 | | | 2007 |
|---|---|---|---|---|
| | США (313) | Японія (216) | Австралія (201) | США (194) |
| Вірусні атаки | 50,132 | 23,286 | 13,356 | 43,256 |
| Неавторизований доступ | 33,920 | 987 | 801 | 5,374 |
| Крадіжки портативних носіїв | 21,222 | 17,450 | 11,279 | 20,005 |
| Витік конфіденційної інформації | 19,277 | 1,066 | 199,632 | 29,304 |
| Відмова в обслуговуванні | 9,335 | 1,195 | 599 | 14,889 |
| Фінансове шахрайство | 8,169 | 231 | 4,684 | 108,890 |
| Інсайдерські зловживання мережевим доступом | 5,909 | 2,685 | 6,517 | 14,895 |
| Мережеве шахрайство | 4,303 | 92 | 2,742 | 3,355 |
| Проникнення в систему ззовні | 242 | 297 | 1,557 | 35,438 |
| Зловживання бездротовою мережею | 1,498 | 52 | - | 2,798 |
| Зловживання при роботі з веб-додатками | 861 | 56 | - | 1,293 |
| Інформаційна диверсія | 830 | 56 | 109 | 5443 |
| Спотворення веб-сайтів | 519 | 178 | 371 | 3,738 |
| Інше | 1,733 | 5,699 | 8,368 | 2,438 |
| Загальні втрати | 167,713 | 53,335 | 250,021 | 345,004 |

Слід мати на увазі, що реальні збитки можуть перевищувати приведені дані, оскільки представники комерційних структур не зацікавлені в оприлюдненні таких результатів.В попередні роки в США близько половини респондентів надавали інформацію про

свої збитки, в 2008 році – лише 28%. Цікаво зазначити, що значну частину збитків (46%) в 2008 році «забезпечили» інсайдери: в межах 1-20% збитків – 25% за рахунок інсайдерів, в межах 21-60% –10%, в межах 61-100% – 11%.

Аналіз збитків, які приносять різні типи електронних атак, в значній мірі визначає інструментарій захисту. В 2008 році майже всі організації використовували антивірусне ПЗ і брандмауери, 85% – віртуальні приватні мережі, 80% - антишпигунське ПЗ. Загалом використання різних типів засобів захисту ілюструють дані табл. 1.4.

Крім прямих збитків від нападів слід врахувати ще витрати на захист інформації. В табл. 1.5 приведено кількість підприємств у відсотках, які витрачають певну частку IT-бюджету на безпеку.

Зазначимо, що витрати на ІБ включають не тільки забезпечення технічними засобами, але й витрати на зовнішніх спеціалістів (зокрема юристів) моніторинг стану ІБ підприємства, аудит, забезпечення аутсорсингу, тощо. Окремою статтею ідуть втрати на освітні і тренувальні програми. В 2009 році в США на такі програми витрачали більше, ніж 10% IT-бюджету – 5% підприємств, 6-10% – 8%, 1-5% – 31%, менше 1% – 42% (13% підприємств не дали такої інформації).

Зазначимо, що витрати на ІБ включають не тільки забезпечення технічними засобами, але й витрати на зовнішніх спеціалістів (зокрема юристів), моніторинг стану ІБ підприємства, аудит, забезпечення аутсорсингу, тощо. Окремою статтею ідуть витрати на освітні і тренувальні програми. В 2009 році в США на такі програми витрачали більше, ніж 10% IT-бюджету – 5% підприємств, 6-10% – 8%, 1-5% – 31%, менше 1% – 42% (13% підприємств не дали такої інформації).

Таблиця 1.4

Використання захисних технологій у відсотках до числа респондентів

| Вид технології | 2006 | | | 2007 | 2008 |
|---|---|---|---|---|---|
| | США | Японія | Австралія | США | США |
| Антивірусне ПЗ | 97 | 94 | 98 | 98 | 97 |
| Міжмережевий екран | 98 | 91 | 98 | 97 | 94 |
| Віртуальні приватні мережі | - | - | - | 84 | 85 |
| Антишпигунське ПЗ | 79 | - | - | 80 | 80 |
| Шифрування даних при передачі | 63 | 32 | - | 66 | 71 |
| Система виявлення вторгнень | 69 | 21 | - | 69 | 69 |
| Системи попередження вторгнень | 43 | 10 | - | 47 | 54 |
| Шифрування збережених даних | 49 | 27 | 43 | 47 | 53 |
| Повторно використовуваний пароль | 46 | 83 | 54 | 51 | 46 |
| Файервол на рівні додатків | 39 | - | - | 45 | 53 |
| Засоби експертизи | 38 | - | - | 40 | 41 |
| Смарт-карти і токени | 38 | - | 24 | 35 | 36 |
| Інфраструктури з відкритим ключем | 36 | - | - | 32 | 36 |
| Централізований контроль доступу | 70 | 75 | 90 | 56 | 50 |
| Захисне ПО на ПК користувачів | 31 | - | - | 27 | 34 |
| Біометричний доступ | 20 | - | 5 | 18 | 23 |
| інше | 4 | 9 | 5 | 4 | 3 |

Підводячи підсумки, зробимо деякі додаткові зауваження. Основна тенденція в сфері кіберкриміналу – потужні, професійніатаки стають більш доступними і прибутковими для злочинців.

Заходи протидії, які застосовують підприємства, (часто після порушення інформаційних систем) спрямовані по таких напрямках:

1) освітні і тренувальні програми – 67% підприємств;
2) додаткові процедури управління і контролю – 58%;
3) розширення сфери застосування криптографії (особливо в портативних засобах) – 58%;
4) удосконалення системи ідентифікації і доступу – 49%;
5) залучення зовнішніх експертів і консультантів – 44%;

6) запобігання втрати даних – 42%;
7) удосконалення захисного програмного забезпечення – 36%.

Таблиця 1.5

Процент втрат на безпеку з бюджету ІТ.

| | 2006 | | | 2007 | 2008 |
|---|---|---|---|---|---|
| | США | Японія | Австралія | США | США |
| Більше 10% | 13% | 8% | 9% | 9% | 15% |
| 8-10% | 10% | 9% | 23% | 11% | 12% |
| 6-7% | 11% | 6% | 23% | 7% | 11% |
| 3-5% | 6% | 18% | 8% | 26% | 8% |
| 1-2% | 26% | 16% | 43% | 23% | 24% |
| Менше 1% | 21% | 16% | 35% | 12% | 18% |
| Не можуть відповісти | 12% | 23% | 16% | 13% | 11% |

Досвід показує, що результати порушень є менш чутливими для підприємств, які мають спеціалізовані підрозділи з ІБ. В 2008 році в США 68% респондентів сповістили, що вони мають такі підрозділи і 18% створюють їх.

Обнадійливим у вирішенні питань збору, систематизації і використання статистичних даних є те, що питання ІБ в різних країнах стають питаннями державної ваги. Зокрема, в США вони віднесені до пріоритетних питань політики президента, і законодавчі органи розробляють і приймають закони, які зобов'язують подавати інформацію про напади на інформаційні структури, а виконавчі органи розробляють стандарти таких реєстрів. Пропонується навіть притягати комерційні структури, які порушують правила ІБ, до цивільної і кримінальної відповідальності.

## ІІ. ЗАДАЧІ І МОДЕЛІ ЕКОНОМІЧНОГО МЕНЕДЖМЕНТУ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

### 2.1 Задачі менеджменту інформаційної безпеки

Розглянемо окремі системи захисту інформації і задачі, які виникають при оптимізації їхніх показників.

**Оптимальний розподіл ресурсів між об'єктами захисту інформації.**Графічно модель цієї задачі представлена на рис. 2.1. Система складається з кількох об'єктів, і задача полягає в пошуку оптимального розподілу ресурсів $\{ y_k^0 \}$ між об'єктами (на нашій моделі для конкретності зображено 3 об'єкти).

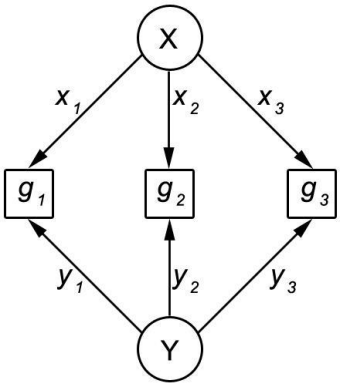



Рис. 2.1 Схема протистояння в інформаційній системі

Задано:

1) розподіл об'ємів інформації $\{ g_k \}$;
2) ресурс захисту $Y = \sum_k y_k$ (ми покладаємо $Y = 1$).

Знайти: розподіл $\{ y_k^0 \}$ при умові, що $X = \sum_k x_k$ і $\{ x_k \}$ невідомі. Ця задача розподіляється на дві частини, в залежності від того, на яку кількість об'єктів здійснюється напад.

**А. Напад здійснюється на один з об'єктів.**Цей варіант реалізується при виконанні принаймні однієї з таких умов:

а) ресурси нападу обмежені, і він вважає недоцільним розпорошувати їх між об'єктами;

б) супернику відомий розподіл інформації $\{ g_k \}$, і він спрямовує свої зусилля на найважливіший об'єкт;

в) суперника цікавить, в першу чергу, інформація на якомусь певному об'єкті, куди він і направляє свої ресурси.

В таких задачах використовується, зазвичай, один з критеріїв пошуку оптимальних рішень в умовах невизначеності [7, 17]. Вибір критерію визначається схильністю до ризику. Нам здається доцільним використати критерій Севіджа, який при розгляді можливих варіантів розподілу $\{x_k\}$ дозволяє знайти розподіл $\{y_k\}$, котрий мінімізує максимальний ризик.

**В. Напад здійснюється на всі об'єкти.** Цей варіант реалізується при виконанні однієї або декількох з таких умов:

а) ресурси нападу достатні для того, щоб розподілити їх між об'єктами;

б) супернику невідомий розподіл об'ємів інформації $\{g_k\}$;

в) суперник виділяє частину ресурсів на розвідку розподілу $\{g_k\}$.

В цьому випадку ми приходимо до задачі лінійного програмування. Застосування симплекс-метода приводить до одного з двох варіантів:

1) розв'язок має сідлову точку, тобто існує в чистих стратегіях; застосовуючи стратегії, кожна сторона одержує найкращий результат, і відхилення від неї може привести до його погіршення;

2) розв'язок існує лише в змішаних стратегіях, які і формують оптимальні рішення.

У випадку двох змінних (тобто двох об'єктів) розв'язок можна одержати графічно, що дає можливість наочно продемонструвати формування оптимального результату. При цьому ми можемо застосувати один з методів умовної оптимізації – метод Лагранжа або метод Якобі [7, 17] і розв'язок одержати аналітично. Якщо кількість змінних перевищує 2, можна застосувати метод оптимізації Белмана [1], який забезпечує найшвидший шлях одержання результату.

В приведених прикладах ми розглядали пряму задачу: задано ресурс захисту $Y$, і необхідно визначити, яким може бути при цьому витік інформації $I$. Може бути сформульована і зворотна задача: задано $I_m$, і необхідно визначити, яким повинен бути ресурс захисту $Y$, котрий забезпечить $I<I_m$. Розв'язок зворотної задачі

утруднений, і тому вона зводиться зазвичай до прямої, а рішення знаходиться методом перебору.

**Багатоцільова задача**. Основними показниками ефективності системи захисту інформації є кількість *I* вилученої інформації і ресурс *Y*, витрачений на її захист. Розглянемо випадок, коли цільова функція включає обидва ці показники з ваговими коефіцієнтами $\lambda$ і $1-\lambda$:

$$s(x,y)=\sum_{k=1}^{l}[\lambda i_k(x,y)+(1-\lambda)y_k],\quad \sum_{k=1}^{l} i_k = i,\quad \sum_{k=1}^{l} y_k = Y.$$

Нашою метою, як і раніше, являється мінімізація цільової функції по *y*, тобто знаходження такого розподілу $\{y_k\}$, при якому досягається мінімум функції *s(x,y)*. Це двокритеріальна задача. В даному випадку крім $g_k$ задаються значення $\lambda$, які можна знайти з умов рівності впливу на економічну безпеку підприємства втрат інформації і витрат на її захист. Необхідно визначити $s_{min}(x,y)$ і відповідний розподіл $\{y_k^0\}$. Можливо залучити третій показник – економічну ефективність $E=\dfrac{\Delta i}{\Delta Y}$, де $\Delta i$ – зменшення втрат інформації при збільшенні витрат $\Delta Y$ на її захист.

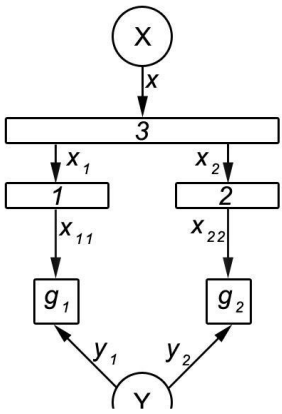


Рис. 2.2 Схема протистояння в багаторубіжній системі

**Багаторубіжний захист.** Комплексні системи захисту інформації зазвичай являються багаторубіжними, або багатоступінчастими. Складність цих систем створює додаткові труднощі як при розрахунку їх показників, так і при визначенні оптимального розподілу ресурсів між окремими рубежами захисту. Спрощена модель багаторубіжної СЗІ показана на рис. 2.2. Слід зазначити, що ресурси захисту розподіляються автономно, тобто ресурси спрямовують на об'єкти $g_1$,$g_2$ і перешкоди 1, 2, 3 паралельно і одночасно. В той же час ресурси нападу направляються на подолання перешкод послідовно, після

подолання чергової перешкоди (рис 2.2). Таким чином, розподіл ресурсів захисту встановлюється заздалегідь, а ресурси нападу розподіляються в динамічному режимі. Послідовний характер подолання перешкод приводить до того, що аналіз протистояння на кожній перешкоді необхідно проводити з врахуванням імовірності подолання попередньої перешкоди. Таким чином, ми одержали стохастичну задачу, в якій повинні бути враховані дві імовірності: імовірність $q(x)$ виділення нападом певних ресурсів $x$ для подолання цієї перешкоди і імовірність $p(x,y)$ подолання перешкоди при певному співвідношенні $x$ і$y$. Маємо умовну імовірність, яка входить в вираз (1) і визначається добутком величин: $q_k(x,y)$, $p_k(x,y)$.

Враховуючи рекурентний характер задачі, одержуємо повну імовірність $P_i$ подолання $i$-ої перешкоди, з врахуванням ймовірностей $p_j$ подолання всіх попередніх перешкод:

$$P_i(x,y)=\prod_{j=1}^{i} p_j(x,y)$$

При цьому кількість величин, які потребують визначення за допомогою експертної оцінки, зростає, що створює додаткові труднощі, проте є особливістю задач пошуку оптимальних рішень в умовах невизначеності.

Зазначимо, що розподіл ресурсів в багаторубіжних системах може бути централізованим – коли кількість об’єктів (і кількість перешкод) невелика, і управління ресурсами ведеться з єдиного центру, і децентралізованим – коли об’єктами являються окремі підприємства, і центр проводить розподіл ресурсів між ними, а подальше управління ресурсами виконується на місцях: $\{y_k\}=\{y_{k,j}\}\sum_j y_{k,j}=y_k$, де $k$ – номер об’єкта (підприємства), $j$ – номер перешкоди. Ступінь

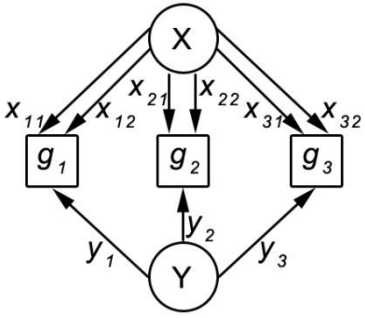


Рис. 2.3 Схема протистояння з використанням розвідки

децентралізації визначається складністю системи. При цьому слід враховувати, що центр має більш повну інформацію, проте окремі підприємства можуть більш оперативно реагувати на дії суперника, хоча з удосконаленням електронних інформаційних систем останнє твердження стає менш значущим.

**Вплив розвідки на розподіл ресурсів**. При наявності достатніх ресурсів напад може свої дії поділити на два етапи – розвідку і здобуття інформації. Схема такого протистояння зображена на рис. 2.3.

Перший індекс в позначенні ресурсів нападу$x_{ks}$– це номер об'єкта, а другий приймає два значення: *s=1*відноситься до розвідки, *s = 2* – до здобуття інформації.

Ця задача відрізняється від задачі 1 більшою кількістю варіантів можливого розподілу ресурсів нападу і тим, що на другому етапі напад діє в умовах певної інформованості, що, звичайно, необхідно враховувати при розподілі ресурсів захисту (цим самим ми переходимо від задачі в умовах невизначеності до задачі в умовах ризику).

Розглянутий варіант можна вважати першим кроком до динамічного управління ресурсами, коли після перших спроб вилучення інформації обидві сторони можуть внести корективи в розподіл своїх ресурсів (для нападу така можливість виникає, якщо спроби виявились вдалими, а для захисту – у випадку, коли спроби суперника зафіксовані).

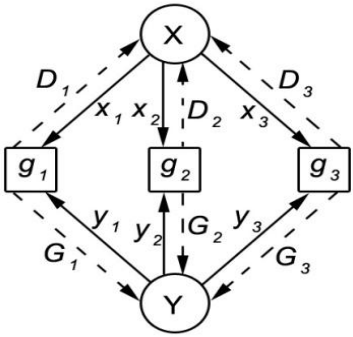



Рис. 2.4 Схема динамічного управління ресурсами

**Управління ресурсами**. Процес динамічного управління ресурсами захисту і нападу схематично показано на рис. 2.4, де для цього використовують сигнали зворотного зв'язку $G_k$ і $D_k$. Приведена схема ілюструє застосування динамічного програмування, яке дає можливість здійснювати оптимальний розподіл ресурсів в динамічному режимі: $x_k=x_k(t)$, $y_k=y_k(t)$. З точки зору теорії ігор це позиційна гра.

**Визначення станів інформаційної безпеки.** В попередніх прикладах ми розглядали результат протистояння двох сторін при спробі вилучення інформації. Не в меншій мірі нас цікавить питання: яким буде стан інформаційної безпеки після низки таких спроб, котрі можуть відрізнятися як параметрами, так і результатом протистояння. Відповідь на це питання може дати використання теорії випадкових процесів, зокрема марковських ланцюгів [18]. При цьому виникає ряд додаткових обчислювальних труднощів, пов'язаних з недостатністю відомостей про характеристики цього випадкового процесу. Зокрема, в розрахунок перехідних ймовірностей, а потім і ймовірностей станів ми повинні закласти:

1) частоту спроб;
2) залежність $q_{kn}(x)$ для $n$-ої спроби;
3) величину $g_{kn}$, яка тепер через зменшення кількості інформації з кожною спробою стає залежною від часу, в результаті чого ланцюги стають неоднорідними.

При розгляді дискретних марковських ланцюгів задача зводиться до розв'язку системи лінійних алгебраїчних рівнянь. Більш повну інформацію можна одержати, розглядаючи неперервні марковські ланцюги, які приводять до необхідності розв'язання системи диференціальних рівнянь Колмогорова.

**Комплексне протистояння**. Ця ситуація виникає, коли кожна сторона захищає свою інформацію і одночасно спрямовує зусилля на здобуття інформації конкурента (звичайно розглядаються тільки легальні методи здобуття інформації). Спрощена схема такого протистояння зображена на рис. 2.5, де через $g$ і $d$ позначені об'єми інформації суперників, верхній індекс 1 відноситься до ресурсів захисту, 2 – до ресурсів здобуття інформації.

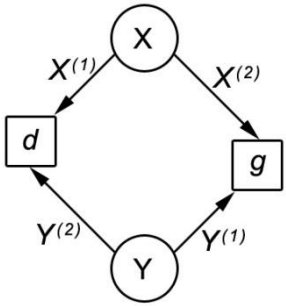


Рис. 2.5 Схема комплексного протистояння

В реальних умовах схема (рис. 2.5), може узагальнювати всі попередні варіанти:

*1)* система містить декілька об'єктів;

*2)* обидві сторони можуть проводити розвідку – і необхідно визначити оптимальну частку ресурсів, які направляються на розвідку і на здобуття інформації (звичайно, з врахуванням можливих дій суперника);

*3)* система захисту кожного об'єкта є багаторубіжною – отже, нам необхідно визначити оптимальний розподіл ресурсів між окремими перешкодами, сформованими по послідовно-паралельній схемі;

*4)* оптимізація ресурсів ведеться в динамічному режимі;

*5)* оптимізація може вестись по декількох критеріях;

Результатом являється визначення станів системи в неперервному режимі.

## 2.2. Проблеми моделювання в сфері інформаційної безпеки

Зростання обсягів і важливості інформації в різних галузях діяльності, розвиток і вдосконалення засобів здобуття і захисту інформації та відповідне збільшення їх вартості привертають все більшу увагу математичного моделювання до ситуацій, які виникають при протистоянні двох сторін у сфері інформаційної безпеки, і пошуку оптимальних рішень поставлених задач.

Складність проблеми обумовлена низкою причин. Головна з них полягає в тому, що пошук оптимального рішення ведеться в умовах невизначеності, коли передбачити дії суперника можна лише з певною імовірністю, а іноді просто неможливо. Серед інших причин відзначимо різноманітність багаторубіжних схемзахисту інформації, складність багатоступінчатої процедурипошуку рішення і різноманітність інструментів – від вибору типу загрози до заключного етапу витоку чи порушення інформації, а також різна природа систем – як фізичних, так і електронних.

Переходячи до систематизації даних, які складають сутність виникаючих проблем, зазначимо, що ці проблеми носять як об'єктивний, так і суб'єктивний характер.

Спробуємо окреслити в загальних рисах головні з них:

1) побудова математичної моделі, яка враховує найбільш важливі аспекти протистояння двох сторін у сфері захисту інформації;

2) вибір критеріїв оптимальності;

3) визначення параметрів розрахунку і функціональних залежностей, які входять в математичну модель;

4) вибір пріоритетів, які враховують важливість для підприємства таких показників, як можливі втрати інформації, витрати на її захист, рентабельність витрат тощо.

Основні задачі, які стоять перед менеджментом:

1) визначення оптимальної кількості ресурсів захисту, мінімізуючих сумарні витрати, які об'єднують потенційні втрати від витоку інформації і витрати на її захист з врахуванням відповідних вагових коефіцієнтів;

2) оптимізація розподілу ресурсів між об'єктами, які містять різні об'єми інформації, характеризуються різним рівнем вразливості і певним ступенем корельованості, а також між окремими рубежами захисту;

3) визначення оптимального розподілу ресурсів в умовах комплексного протистояння в конкурентній боротьбі, коли кожна сторона захищає свою інформацію і одночасно спрямовує свої зусилля на здобуття інформації конкурента, причому частина ресурсів може бути направлена на розвідку;

4) визначення зміни станів інформаційної безпеки з часом із врахуванням можливих дій суперників;

5) розробка методики управління ресурсами в динамічному режимі, в якій враховані приведені ситуації та показники.

Перший крок у дослідженні процесів протистояння – це розробка аналітичної математичної моделі. Побудову моделі можна поділити на декілька етапів.

I. Визначення показників системи захисту інформації:

1) кількість $l$ об'єктів захисту;
2) об'єм $g_k$ інформації на кожному об'єкті ($k$ – номер об'єкту);
3) виділений ресурс $Y$ захисту;
4) початкова вразливість $\upsilon_{k0}$ об'єкта, яка обумовлена його природною захищеністю і визначається як імовірність результативної атаки при $Y=0$;
5) відношення до ризику.

II. Оцінка дій суперника:

1) характер атак (націлені, ненацілені);

2) виділений ресурс $X$ нападу;
3) імовірність $p_k$ нападу на об'єкт;
4) імовірність виділення ресурсів $x_k$.

II. Формування цільової функції, яке включає вибір цільового показника і незалежної змінної та встановлення між ними функціональної залежності:

1) цільовий показник:
   а) кількість вилученої інформації;
   б) сумарні витрати ресурсів, які включають втрати від вилучення інформації і витрати на її захист;
   в) ефективність інвестування, яку визначаємо як частку двох величин – зменшення об'єму вилученої інформації і витрат на захист;
2) незалежні змінні:
   а) ресурси нападу і захисту – $x_k$ і $y_k$;
   б) динамічна вразливість $\upsilon_k(x,y)$;
3) вид функціональної залежності цільового показника від незалежної змінної:
   а) степенева;
   б) показникова.

Проведемо перелік операцій, які необхідно здійснити, приступаючи до рішення задачі (інакше кажучи, складемо мережевий графік).

Отже, нам необхідно:

1) сформулювати проблему дослідження (в нашому випадку – визначити оптимальний розподіл ресурсів між об'єктами);

2) зібрати інформацію про характеристики системи і очікувані умови протистояння (кількість інформації на об'єктах, їх вразливість, імовірність нападу на кожен з об'єктів, імовірність виділення певної кількості ресурсів нападу на кожен з об'єктів);

3) побудувати математичну модель (сформувати цільову функцію $i$, встановити значення параметрів і форми залежностей, які входять в цю функцію);

4) вибрати метод розв'язку задачі (аналітичний, геометричний);

5) скласти і налагодити програму для комп'ютера;

6) провести розрахунки і представити результати в найбільш лаконічній, доступній і інформативній формі (таблиці, графіки, рисунки);

7) представити висновки і рекомендації (оптимальний розмір і розподіл ресурсів при різних умовах протистояння).

### 2.3 Моделі інформаційної безпеки

Розробка перших моделей антагоністичного протистояння була викликана потребами планування військових операцій. Найбільше відображення військове протистояння знайшло в задачі Гроса [14], в якій дві конфліктуючі сторони володіють ресурсами *X* і*Y*, розподіленими між окремими об'єктами, а цільова функція визначає необхідну кількість засобів нападу, які можуть прорватися через оборону. В застосуванні до задач інформаційної безпеки *i(x,y)*виражає потенціальні втрати інформації в залежності від ресурсів нападу і захисту – *x* і, відповідно, *y*– і має вигляд:

$$i_{ij}(x, y) = \sum_{k=1}^{l} g_k \, p_{ijk} \, f(x_{ik} - y_{jk}), \qquad (2.1)$$

де *i* та *j*– номери варіантів розподілу ресурсів;

*k* – номер об'єкту;

$p_{ijk}$ – імовірність *ij*-го розподілу ресурсів на *k* -му об'єкті,

$$f(x_{ik} - x_{jk}) = \begin{cases} 0 & при \quad x_{ik} - x_{jk} \le 0 \\ x_{ik} - x_{jk} & при \; 0 < x_{ik} - x_{jk} < 1 \\ 1 & при \quad x_{ik} - x_{jk} \ge 1 \end{cases}$$

Найбільш відомою моделлю, в якій розглядаються економічні питання протистояння в інформаційній сфері, є модель Гордона-Лоеба (ГЛ) [22]. В цій моделі цільова функція визначає зменшення втрат від вилучення інформації за рахунок внесення інвестицій з відрахуванням витрат *y* на її захист (в [22] ці витрати позначені через *z*). В функцію у вигляді параметра входить вразливість *υ* об'єкта, котра визначається як імовірність того, що напад буде успішним при *y*=0. Авторами [22] запропоновано два широких

класи функцій $S(y,\upsilon)$, котрі визначають імовірність порушення інформації:

$$S^{I}(y;v)=\frac{v}{(\alpha y+1)^{\beta}}; \qquad (2.2)$$

$$S^{II}(y;v)=v^{\alpha\cdot y+1}, \qquad (2.3)$$

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

$$E(y)=[\upsilon - S(y,\upsilon)]\,L - y, \qquad (2.4)$$

де $L$ – потенційні втрати інформації при здійсненні нападу.

В (2.4) перша складова ($\upsilon L$) визначає кошти, втрачені в результаті нападу при відсутності системи захисту інформації (СЗІ), друга ($S(y,\upsilon)\cdot L$) – при введенні СЗІ, третя ($y$) – інвестиції в СЗІ. Загалом (3) визначає кошти, збережені за рахунок введення СЗІ.

Метою аналізу в [22] є визначення оптимальних витрат $y^0$ при різних значеннях вразливості. Показником оптимальності є максимум прибутку від інвестицій, що виражається умовою $E'(y)=0$.Показано, що вид залежності $y^0(\upsilon)$ відрізняється для двох класів функцій $S(y,\upsilon)$, і для розробки рекомендацій по визначенню раціональної кількості інвестицій, крім вибору виду функції $S(y,\upsilon)$, необхідно встановити рівень вразливості об'єкта.

Останнім часом з'явилась низка робіт [23-25], направлених на розвиток моделі ГЛ. Зокрема, в [24] 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 виділити три варіанти їх взаємодії: відсутність впливу, позитивний вплив і негативний вплив. Для більш детального дослідження цього питання в [24] введені поняття продуктивності зменшення вразливості і продуктивності зменшення загрози, а також простору продуктивності, який об'єднує ці показники. В залежності від їх значень випливають висновки відносно вибору стратегій, які забезпечують оптимальні витрати на захист інформації.

Слід зазначити, що оптимізація розподілу ресурсів між об'єктами в значній мірі зумовлена можливими стратегіями нападу. Напади суперника можуть бути не націленими (шкідливе програмне забезпечення, віруси, фішинг, спам тощо) і націленими (наприклад, хакерські атаки на банківську базу даних з метою вилучення коштів). Характер нападів може бути обумовлений, в одних випадках, – цільовою спрямованістю зловмисника, в інших, – кількістю ресурсів конкурента. Останній варіант спостерігається, зокрема, коли об'єкти однотипні, і конкурент розглядає доцільність розподілу обмежених ресурсів між окремими об'єктами. Націлені атаки трапляються рідше, ніж не націлені, проте їх наслідки можуть бути більш серйозними для підприємства, що, звичайно, слід враховувати при розподілі ресурсів захисту. В [23] зроблено припущення, що клас функцій (2.2) краще описує націлені атаки, а клас (2.3) – не націлені.

В [13] запропоновано інший підхід до поставленої проблеми. Математична модель [13] передбачає використання цільової функції $i(x,y)$, де $i$ – відносна кількість вилученої інформації, $x$ і $y$ – ресурси нападу і, відповідно, захисту. Ця функція в загальних рисах має вигляд:

$$i(x,y)=\sum_{k=1}^{l} i_k(x,y)=\sum_{k=1}^{l} g_k \cdot p_k \cdot q_k(x,y)\cdot f_k(x,y), \qquad (4)$$

де $k=\overline{1,l}$ – номер об'єкта;

$g_k$– відносна кількість інформації на об'єкті;

$p_k$ – імовірність нападу на об'єкт;

$q_k(x,y)$ – імовірність виділення нападом ресурсів $x$ на $k$-ий об'єкт ($y$в цю функцію входить як параметр);

$f_k(x,y)$ – залежність частки вилученої інформації від співвідношення $x$ та $y$, яку можна розглядати як імовірність вилучення інформації при заданих значеннях $x$ та $y$.

Основна вимога до залежностей$f(x,y)$:при$\frac{x}{y}\to 0 \quad f(x,y)\to 0,$

при $\frac{x}{y}\to\infty \quad f(x,y)\to a,$де $a \le 1$ - максимально можлива кількість вилученої інформації, яка визначається специфікою

об'єкта і його системи захисту. Цим умовам відповідають степеневі і показникові функції:

$$f(x,y)=\frac{a(x/y)^n}{(x/y)^n+c} \qquad f(x,y)=a(1-e^{-m(x/y)^n})$$

Враховуючи, що при відповідному виборі параметрів ці залежності можуть стати досить близькими, обмежимось розглядом степеневих функцій. Параметри *n* та *c* в степеневій функції *f(x,y)* можна встановити, виходячи з наступних міркувань.

1. При $\frac{x}{y}>>1$ залежність *f(x,y)* при різних *n* та *c* повинні мати схожий характер, який диктується умовою $f(x,y)\to a$. При $\frac{x}{y}\approx 0$ опуклість кривої може бути направлена як вгору, так і вниз–в залежності від початкової вразливості об'єкта. В математичному виразі степеневої залежності *f(x,y)* опуклість направлена вгору при $n\le 1$, а при $n>1$ –вниз.

2. Вважаючи, що втрата 10-15% інформації є для підприємства досить відчутною, а 15-20%–критичною, формулюємо умову: при $\frac{x}{y}\approx 1 \quad f(x,y)\approx 0{,}05..0{,}15$, при $\frac{x}{y}\approx b \quad f(x,y)\approx 0{,}2..0{,}3$.

Граничне значення $\frac{x}{y}\approx b$ обирається з наступних міркувань. Кількість ресурсів, які можуть бути виділені на захист інформації, за статистичними оцінками становить $y\approx 0..0{,}15g$. Вважаємо, що витрати ресурсів сторони нападу лежать в інтервалі $x\approx 0..0{,}5g$ (подальше їх збільшення визнаємо недоцільним). Виходячи з реальних граничних витрат та вважаючи, що витрати обох сторін визначаються одними і тими ж показниками (важливістю об'єкта та його вразливістю) і тому змінюються синхронно, отримуємо граничне значення $\frac{x}{y}=\frac{0{,}5}{0{,}15}\approx 3$. В подальшому ресурси нападу будемо подавати у відносних величинах в двох варіантах: віднесені

до кількості інформації– $\frac{x_k}{g_k}$, $x_k = 0..1$ або до ресурсів захисту– $\frac{x_k}{y_k} = 0..3$. Значення $\frac{x}{y}$, які лежать за межами останнього інтервалу (зокрема, при $y \to 0$ $\frac{x}{y} \to \infty$) можуть бути розглянуті окремо. Варто зазначити, що приведені величини є орієнтовними і в окремих випадках можуть бути перевищені.

3. При невеликих значеннях $\frac{x}{y}$ $\left(\frac{x}{y} \le 1\right)$ СЗІ повинна бути рентабельною, тобто зменшення втрат інформації *ΔI* повинно перевищувати витрати *y* на її захист. При значних величинах $\frac{x}{y} \approx 3$ рентабельним повинен бути напад: *i≥x.*

Положення кривих *f(x)* (для спрощення прийнято*y=1*) показано на рис. 2.6.

1. $f(x) = \frac{x}{x+1}$ 2. $f(x) = \frac{x^2}{x^2+4^2}$ 3. $f(x) = \frac{x^4}{x^4+4}$

4. $f(x) = \frac{x^4}{x^4+2^4}$ 5. $f(x) = 1 - e^{-\frac{1}{2}x}$ 6. $f(x) = 1 - e^{-\frac{1}{2}x^2}$

7. $f(x) = 1 - e^{-x^2}$ 8. $f(x) = 1 - e^{-\frac{1}{4}x^4}$ 9. $f(x) = 1 - e^{-\frac{1}{2}x^4}$

Друга серйозна задача – це визначення виду функції*q(x)*. Найпростіший варіант дає лапласівський підхід, відповідно до якого всі можливі рішення суперника в умовах невизначеності вважаються рівноймовірними, тобто*q(x)=const*. Звичайно, такий підхід потребує серйозного коригування, оскільки вимагає визначення інтервалу *Δx*, в якому знаходяться можливі значення *x,* а головне – врахування того факту, що в реальних умовах *q(x)≠const* Таким чином, виникають два питання: 1) визначення виду залежності *q(x);* 2) встановлення значення $x_m$, при якому залежність *q(x)* досягає максимуму. За нашою думкою, ця

залежність має більшу крутизну на ділянці $x=0..x_m$, поступово спадаючи при зростанні $x$ і прямуючи до нуля при $x>>1$. Цим умовам задовольняють залежності виду $f(x)=Nx^n e^{-h^2x^2}$, зокрема розподіл Максвела $q_M(x)=Nx^2e^{-h^2x^2}$ і розподіл Релея $q_P(x)=Nxe^{-h^2x^2}$, де $N$ – нормуючий коефіцієнт, а параметри $n$, $h$ визначають положення максимуму залежності (для розподілу Максвела $x_m=\frac{1}{h}$, розподілу Релея $x_m=\frac{1}{\sqrt{2}h}$) і ступінь її асиметрії.

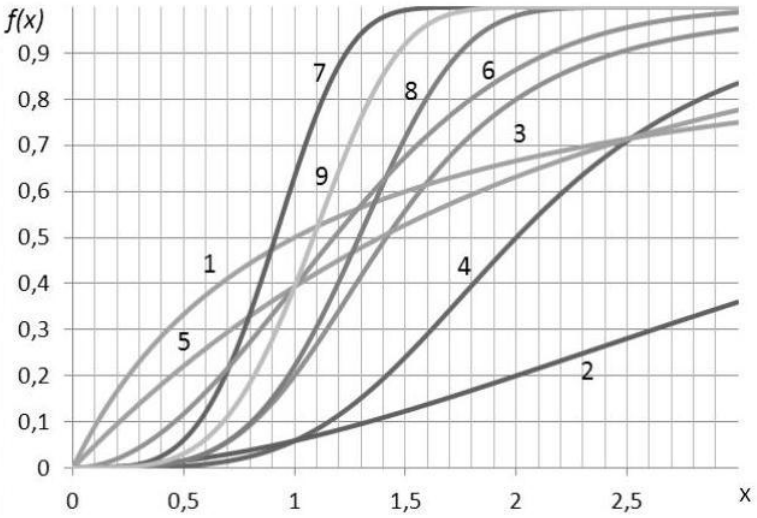



Рис. 2.6. Залежність втрат інформації від співвідношення ресурсів нападу і захисту.

На рис. 2.7. приведені максвеловські розподіли для трьох значень $x_m$, а також лінія $q(x)=q=const$, проведена в межах $x=0..3$, де величина $q$ визначається умовою, що повна імовірність події, тобто площа під відповідними залежностями, дорівнює 1. Штриховими лініями зображені релеєвські розподіли. Зазначимо, що основна відмінність форм розподілів Максвела і Релея, суттєва для наших задач, полягає в тому, що в початковій області (при $x \approx 0$) опуклість

в розподілі Максвела направлена вниз, а в розподілі Релея – вгору (рис. 2.7). Зрозуміло, що комбінація можливих залежностей *f(x)* і *q(x)* утворює значну кількість варіантів, яку бажано обмежити. Це можна зробити, враховуючи специфіку кожної системи захисту інформації і спираючись на оцінку експертів.

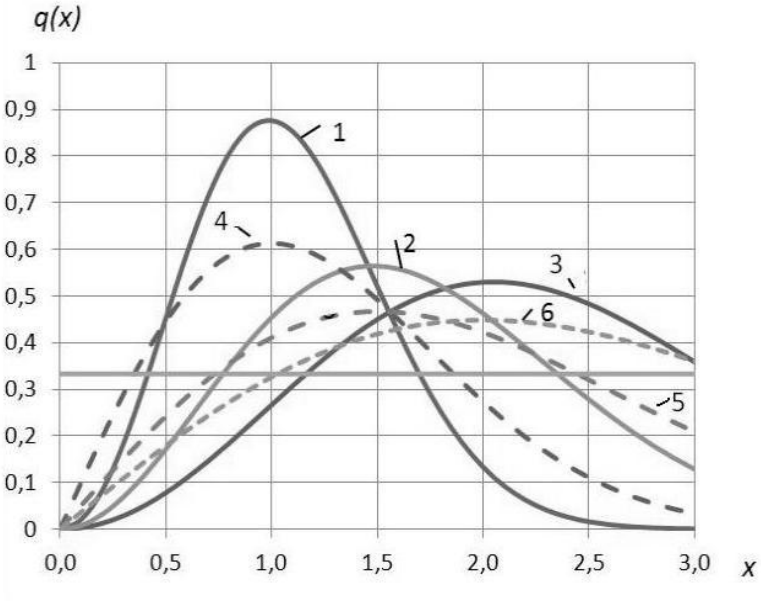



Рис. 2.7. Диференціальна функція розподілу відносних ресурсів нападу.

1,2,3 – розподіл Максвела, 4,5,6 – розподіл Релея.

1)*h=1; N=2,257;* 4) *h=0,7071; N=1,011;*

2)*h=0,66; N=0,682;* 5) *h=0,471; N=0,513;*

3) *h=0,5; N=0,358;* 6) *h=0,3536; N=0,37;*

Коефіцієнти*h*обрані таким чином, що максимуми функцій розподілу знаходяться в точках*x=1* (криві 1, 4); *x=1,5* (криві 2, 5); *x=2* (криві 3, 6). Коефіцієнти N нормують інтеграл функції в межах [0;3] до одиниці.

Розглянемо тепер обчислювальні аспекти поставленої задачі. Після того, як встановлено вид залежностей*f(x,y), q(x,y),* можна переходити до розрахунку значень цільової функції (4). Зазначимо, що рішення ускладнюється значною кількістю змінних величин,

котрі впливають на кінцевий результат. Ці змінні можна поділити на дві групи.

I. Змінні параметри, які задаються стороною захисту (керовані змінні):

1) ресурс $Y$ захисту;
2) розподіл ресурсів $y_k$ по об'єктах захисту $\sum_{k=1}^{l} y_k = Y$;
3) значення вагових коефіцієнтів $g_k$, які характеризують об'єм інформації на кожному з об'єктів з врахуванням її важливості;
4) допустимий рівень ризику $R$ втрати інформації, який визначає максимально допустиме значення $i$.

II. Невизначені параметри, які задаються стороною нападу (некеровані змінні):

1) ресурс $X$ нападу;
2) розподіл ресурсів нападу $x_k$ по об'єктах, $\sum_{k=1}^{l} x_k = X$;
3) імовірності $p_k$ нападу на кожний з об'єктів.

Рішення знаходиться як оптимум цільової функції по одному з відомих критеріїв, причому при постановці задачі можливі два підходи:

1) однокритеріальний підхід, при якому оптимізації підлягає одна з характерних величин – кількість вилученої інформації $i(x,y)$, ресурс захисту $Y$ чи рівень ризику $R(Y)$;

2) багатокритеріальний підхід – коли оптимум шукають одночасно по декількох критеріях, при цьому, враховуючи неможливість (у багатьох випадках) одночасного досягнення оптимуму по окремих критеріях, встановлюють їх ієрархію або додаткові обмеження на кожен з них.

В залежності від нашої інформованості про дії суперника будемо розглядати дві ситуації. Якщо імовірність настання окремих подій, які впливають на кінцевий результат, може бути встановлена з деякою мірою точності, то прийняття рішень в термінах ризикології відбувається „в умовах ризику", якщо ж така імовірність через брак інформації не може бути встановлена, то „в умовах невизначеності".

Є декілька шляхів визначення величин, які входять в розрахунок. Перший – це використання статистичних даних. Таким чином можна встановити значення такого параметра, як імовірність нападу. Другий – експертна оцінка, за допомогою якої можна визначити, наприклад, відносну важливість об'єктів, тобто значення параметрів $g_k$. Величини $x_k$ і $y_k$ утворюють дві групи величин, одна з яких фіксується і використовується в розрахунку як параметр, а друга є незалежною змінною (цей вибір залежить від того, для якої сторони – нападу чи захисту – ставиться задача).

## Контрольні питання

1. Типи задач, які виникають при аналізі протистояння в сфері інформаційної безпеки.
2. Проблеми, які виникають при побудові математичної моделі інформаційного протистояння.
3. Модель Гроса в застосуванні до задач інформаційної безпеки.
4. Модель Гордона-Лоеба і її модифікації.
5. Форма цільової функції, яка визначає кількість вилученої інформації, аналіз її складових.
6. Форми функцій, які визначають частку вилученої інформації в залежності від ресурсів нападу і захисту.
7. Диференціальна функція розподілу ресурсів нападу, її форми.
8. Обчислювальні аспекти розрахунку значень цільової функції.

## ІІІ. ПОКАЗНИКИ СИСТЕМ ЗАХИСТУ ІНФОРМАЦІЇ І МЕТОДИ ЇХ ОЦІНКИ

### 3.1 Показники багаторубіжних систем захисту інформації

Системи захисту інформації (СЗІ) являються звичайно багаторубіжними, або багатоступінчастими (периметр, приміщення, інтер'єр, тощо). Розрахунок їх показників є важливою задачею, розв'язок якої ускладнюється тим, що формулювання деяких з них не є загальновизнаними, а інші мають різні трактування. Одним з таких показників є надійність. Специфіка розрахунку цього показника обумовлена тим, що функції елементів звичайних, зокрема, радіоелектронних систем відрізняються від функцій СЗІ. В першому випадку елементи системи призначені для пропускання сигналу, в другому – для його блокування (під сигналом в СЗІ розуміємо будь-які спроби проникнення в об'єкт захисту). Тому й вигляд виразів для розрахунку надійності при послідовному і паралельному з'єднанні елементів СЗІ порівняно з радіоелектронними системами змінюється на протилежний.

Відповідно до цього надійність *R*радіоелектронної системи визначає імовірність пропускання сигналу всією системою, а надійності $p_i$ , $q_i$ – імовірності пропускання сигналу елементами системи при їх послідовному і паралельному з'єднанні (рис.3.1).

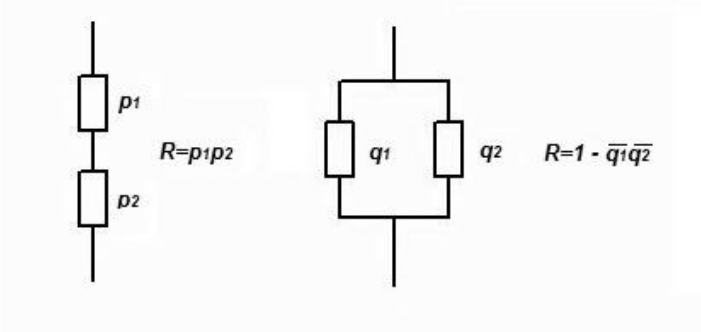



Рис. 3.1 Надійність радіоелектронної системи при послідовному і паралельному з'єднанні елементів.

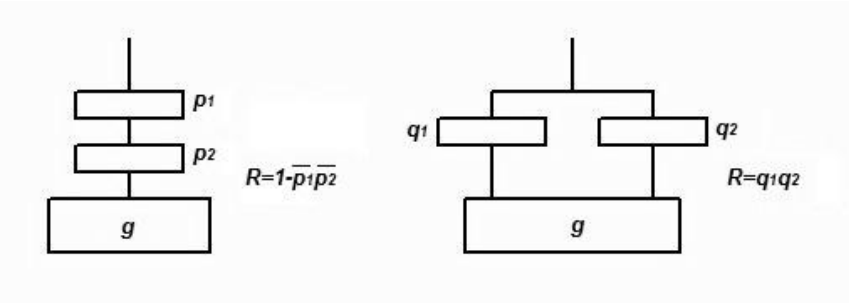



Рис.3.2 Надійність СЗІ при послідовному і паралельному розташування перешкод.

В СЗІ величина*R* визначає імовірність того, що проникнення в систему не відбудеться, а$p_i$ і $q_i$ – імовірності блокування сигналу її елементами (рис. 3.2).

Розглянемо об'єкт, який містить інформацію в двох різних формах об'ємами $g_1$ і, відповідно, $g_2, g_1 + g_2 = g$ (це можуть бути також два об'єкти або два канали), і СЗІ, котра може бути сформована за різними схемами з'єднання елементів – паралельного, послідовного і послідовно – паралельного (рис.3.3).

В схемах на рис. 3.3,а,б СЗІ захищає одночасно обидва об'єми інформації – $g_1$ і $g_2$, в схемі на рис.3.3,в система складається з двох підсистем, кожна з яких захищає один з об'ємів – $g_1$ або $g_2$, схема на рис. 3.3,г поєднує обидва варіанти.

Візьмемо за базову схему рис. 5.3а і порівняємо з нею інші схеми (рис. 3.3,б,в,г) по основних показниках – таких, як надійність *R*, кількість $i$ вилученої інформації, її вартість *Y*, ефективність використання ресурсів захисту *E*, котру визначимо як частку двох величин - зменшення втрат інформації в результаті застосування СЗІ і витрат на захист інформації:

$$E = \frac{i_0 - i}{Y},$$

де $i_0$ і $i$ - можливі втрати інформації при відсутності СЗІ, тобто при відсутності перешкод, і при їх наявності. Нижнім нуликом позначені величини, які характеризують об'єкт при відсутності

СЗІ, верхнім нуликом – величини, які відносяться до базової моделі, зображеної на рис.3.3,а. Вирази для цих показників приведені на рис. 3.3. Величину $i$ будемо розраховувати як математичне сподівання втрат, величини $g$ і $i_0$ вважаються відомими.

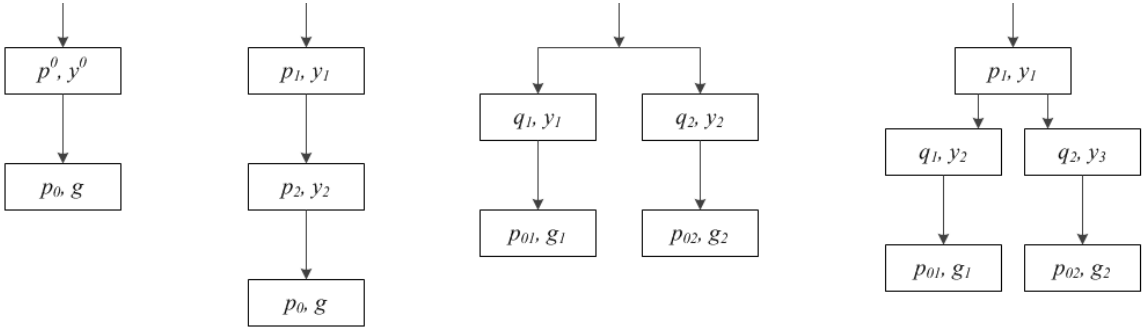
Ускладнення СЗІ (перехід від схеми на рис. 3.3,а до наступних схем) дозволяє покращити деякі характеристики за рахунок збільшення вартості. Співставлення зміни цих показників, яке визначає доцільність ускладнення схеми, дає результат, який залежить від параметрів надійності ( $p_i, q_j$ ) нових елементів схеми і їх вартості ( $y_i, y_j$ ).

Проілюструємо це на числовому прикладі. Задамо параметри базової моделі (рис.3.3,а), подаючи всі величини у нормованій формі:

$g = 1$; природна захищеність $p_0 = 0{,}6$; $p^0 = 0{,}9$; $y^0 = 0{,}1$, звідки $R^0 = p^0 = 0{,}9$; $i^0 = \overline{p^0}\,\overline{p_0} g = 0{,}4 \cdot 0{,}1 = 0{,}04$;

$$E = \frac{\overline{p_0}(1-\overline{p^0})g}{Y} = \frac{0{,}4(1-0{,}1)\cdot 1}{0{,}1} = 3{,}6.$$

Перехід до схеми на рис. 3.3,б при використанні таких же елементів захисту, що і в схемі на рис.3.3а ( $p_1 = p_2 = p^0$, $y_1 = y_2 = y^0$ ) при збільшенні вартості вдвічі дає змогу збільшити надійність з $R^0 = 0{,}9$ до $R = 0{,}99$ та зменшити імовірний витік інформації з $i^0 = 0{,}04$ до $i = 0{,}01$ при зменшенні ефективності з $E^0 = 3{,}6$ до $E = 1{,}98$. Зменшення ефективності пов'язане з тим, що вартість СЗІ може збільшуватись суттєво (в декілька разів), в той же час, як величина витоку інформації обмежена значенням $g$ і зі зменшенням втрат при ускладненні СЗІ прямує до нуля. Тому питання про доцільність ускладнення схеми вирішується з огляду на виділені ресурси захисту і пріоритети її показників.




а)

$$R^0 = p^0$$
$$i^0 = \overline{p^0}\,\overline{p_0}g$$
$$Y^0 = y^0$$
$$E^0 = \frac{\overline{p^0}(1-\overline{p^0})g}{Y^0}$$

б)

$$R = 1 - \overline{p_1}\,\overline{p_2}$$
$$i = \overline{p_1}\,\overline{p_2}\,\overline{p_0}g$$
$$Y = y_1 + y_2$$
$$E = \frac{\overline{p_0}(1-\overline{p_1}\,\overline{p_2})g}{Y}$$

в)

$$R = q_1 q_2$$
$$i = \overline{q_1}\,\overline{p_{01}}g_1 + \overline{q_2}\,\overline{p_{02}}g_2$$
$$Y = y_1 + y_2$$
$$E = \frac{\overline{p_{01}}(1-q_1)g_1 + \overline{p_{02}}(1-q_2)g_2}{Y}$$

г)

$$R = 1 - \overline{p_1}(1-q_1q_2)$$
$$i = \overline{p_1}(\overline{q_1}\,\overline{p_{01}}g_1 + \overline{q_2}\,\overline{p_{02}}g_2)$$
$$Y = y_1 + y_2 + y_3$$
$$E = \frac{\overline{p_{01}}(1-p_1q_1)g_1 + p_{02}(1-p_1q_2)g_2}{Y}$$

Рис. 3.3 Схеми з'єднання елементів СЗІ і їх показники

Для схеми рис.3.3,в при використанні тих же елементів захисту ($q_1 = q_2 = p^0$, $y_1 = y_2 = y^0$) і вважаючи для спрощення розрахунків розподіл об'ємів інформації рівномірним ($g_1 = g_2 = \frac{g}{2}$), одержуємо: $R = 0{,}81$; $Y = 0{,}2$; $i = 0{,}1$; $E = 1{,}5$. Порівнюючи цю схему з базовою, бачимо, що при однакових втратах інформації ($i = i^0$) схема на рис.3.3,в потребує вдвічі більше ресурсів, що приводить до зменшення ефективності в 2 рази. Зменшення надійності ($R = 0{,}81$ порівнюючи з $R^0 = 0{,}9$) можна пояснити тим, що схема на рис.3.3,в зображає фактично два канали захисту інформації і, оскільки при розрахунку надійності ми вважаємо, що перешкоди повинні спрацьовувати в обох каналах, то це й приводить до зменшення результуючої імовірності. Зазначимо, що надійність радіоелектронної системи по схемі на рис.3.3,в буде визначатись умовою, що достатньо функціонування принаймні одного каналу для пропускання сигналу, в той же час, як і в СЗІ така ситуація являється неприйнятною, оскільки при цьому буде спостерігатись витік інформації по іншому каналу.

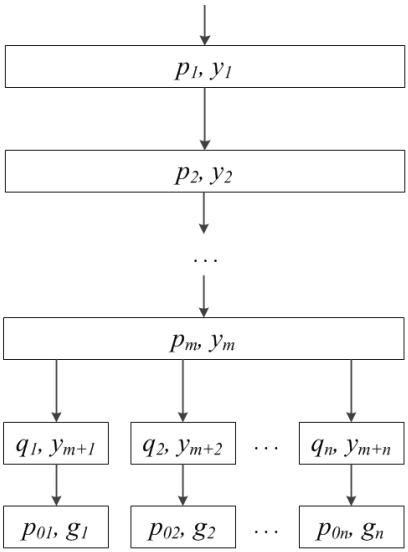
Перехід до схеми на рис.3.3,г з параметрами елементів $q_1 = q_2 = q_3 = p^0$, $y_1 = y_2 = y_3 = y^0$ дає такі показники: $R = 0{,}98$; $Y = 0{,}3$; $i = 0{,}01$; $E = 1{,}33$. Як бачимо, схема з рис. 3.3г порівняно зі схемою на рис.3.3,в має більшу надійність і меншу імовірність витоку інформації, але в результаті збільшення вартості − меншу ефективність. Одержане значення $E = 1{,}33$ показує, що при заданих параметрах елементів СЗІ надмірне ускладнення схеми недоцільне: при розподілі об'єму інформації на $g_1$ і $g_2$ краще використовувати схему рис.3.3,в або застосовувати елементи схеми з кращими параметрами – більшими значеннями $p$ і $q$, меншою вартістю $y$. Розрахункові вирази для СЗІ в узагальненій формі приведені на рис. 3.4.

Поставимо питання: що раціональніше – вкласти всі ресурси захисту в одну перешкоду, чи розподілити їх між декількома перешкодами, розташованими послідовно. Для цього розглянемо три варіанти схеми (рис. 3.3), які відрізняються кількістю

перешкод: 1-ий – одна перешкода, 2-ий – дві перешкоди, 3-ій – три перешкоди. Порівняємо ці варіанти за умови, що ресурси захисту для трьох схем однакові і становлять $Y = 0{,}1$. Вважатимемо, що ресурси, виділені на $i$-ту перешкоду у випадку декількох перешкод, розподілені рівномірно:

$$y_i = \frac{Y}{m},$$

де $m$ - кількість перешкод.




$$R = 1 - \prod_{i=1}^{m} \overline{p_i} \left(1 - \prod_{j=1}^{n} q_j\right)$$

$$i = \prod_{i=1}^{m} \overline{p_i} \cdot \sum_{j=1}^{n} q_j p_{0j} g_j$$

$$Y = \sum_{i=1}^{m} y_i + \sum_{j=1}^{n} y_j$$

$$E = \frac{\sum_{j=1}^{n} p_{0j} \cdot \left(1 - \prod_{i=1}^{m} \overline{p_i} q_i\right) g_j}{Y}$$

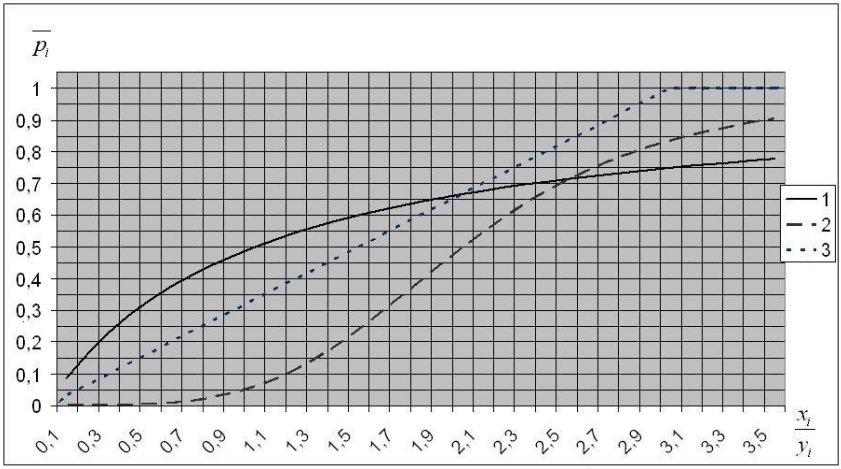
Рис. 3.4 Загальний вигляд схеми багаторубіжної СЗІ.

Використаємо залежність імовірності подолання перешкоди від величини $\frac{x_i}{y_i}$ у вигляді (рис. 3.5):

$$\overline{p_i} = \frac{x_i / y_i}{1 + x_i / y_i},$$

де $x_i$ - кількість ресурсів нападу, виділених на подолання $i$-ї перешкоди, $y_i$ - кількість ресурсів, виділених на її захист.

Вважаємо, що на подолання першої перешкоди з виділених $X = \sum_{i=1}^{m} x_i$ ресурсів напад витрачає $x_1 = y_1$ ресурсів. Тоді на подолання наступної перешкоди у нападника залишається $x_2 = X - x_1$ ресурсів, а на подолання *i*-ї : $x_i = X - \sum_{\alpha=1}^{\alpha-1} x_\alpha$ .




$$1 - \overline{p_i} = \frac{x_i / y_i}{1 + x_i / y_i} \qquad 2 - \overline{p_i} = \frac{\left(x_i / y_i\right)^4}{\left(x_i / y_i\right)^4 + 2^4} \qquad 3 - \overline{p_i} = \begin{cases} \frac{1}{3} \cdot \frac{x_i}{y_i} & при \quad \frac{x}{y} \le 3 \\ 1 & при \quad \frac{x}{y} \ge 3 \end{cases}$$

Рис. 3.5. Залежність імовірності подолання перешкоди від відношення $\frac{x}{y}$

Розрахуємо показники надійності $R = 1 - \prod_{i=1}^{m} \overline{p_i}$, кількості вилученої інформації $i = \prod_{i=1}^{m} \overline{p_i} \cdot g$, ефективності використання

ресурсів захисту $E = \frac{i_0 - i}{Y}$ для кожної схеми та визначимо цільову функцію, яка характеризує ефективність всієї системи:

$$W = k_1 R + k_2 Y + k_3 i + k_4 E \to max$$

Вагові коефіцієнти $k_1, k_2, k_3, k_4$ визначаються методом експертної оцінки. Розглянемо два випадки вибору коефіцієнтів $k$. В першому основну вагу надамо технічним показникам $R$ і $i$ (при цьому найбільше значення мають коефіцієнти $k_1$ і $k_3$).

Цільову функцію в цьому випадку приймемо у вигляді:

$$W_1 = 0{,}4 \cdot R - 0{,}1 \cdot Y - 0{,}3 \cdot i + 0{,}2 \cdot E \to max$$

В другому випадку основну вагу надамо економічним показникам $Y$ і $E$, і цільову функцію оберемо у вигляді:

$$W_2 = 0{,}2 \cdot R - 0{,}3 \cdot Y - 0{,}1 \cdot i + 0{,}4 \cdot E \to max$$

Жирним шрифтом виділені оптимальні значення показників. За результатами, поданими в таблиці 3.1, можна зробити висновок, що в умовах даної задачі цільова функція набуває максимального значення при застосуванні трьох перешкод, причому дана схема виявляється найбільш ефективною за всіма показниками.

Таблиця 3.1

Результати розрахунків показників СЗІ при використанні залежності $\overline{p_i}(x, y)$ у вигляді 1 (рис. 3.5)

| Показники | Значення показників залежно від кількості ресурсів нападу та кількості перешкод | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $X = 0{,}1$ | | | $X = 0{,}2$ | | | $X = 0{,}3$ | | |
| | Схема 1 | Схема 2 | Схема 3 | Схема 1 | Схема 2 | Схема 3 | Схема 1 | Схема 2 | Схема 3 |
| Вартість перешкоди, $y$ | 1 | 0,5 | 0,33 | 1 | 0,5 | 0,33 | 1 | 0,5 | 0,33 |
| Імовірність подолання $i$-ї перешкоди, $\overline{p_i}$ | $\overline{p_1}$=0,5 | $\overline{p_1}$=0,67<br>$\overline{p_2}$=0,5 | $\overline{p_1}$=0,75<br>$\overline{p_2}$=0,67<br>$\overline{p_3}$=0,5 | $\overline{p_1}$=0,67 | $\overline{p_1}$=0,8<br>$\overline{p_2}$=0,75 | $\overline{p_1}$=0,857<br>$\overline{p_2}$=0,83<br>$\overline{p_3}$=0,8 | $\overline{p_1}$=0,75 | $\overline{p_1}$=0,857<br>$\overline{p_2}$=0,83 | $\overline{p_1}$=0,9<br>$\overline{p_2}$=0,89<br>$\overline{p_3}$=0,875 |
| Надійність захисту, $R$ | 0,5 | 0,667 | **0,75** | 0,333 | 0,4 | **0,429** | 0,25 | 0,286 | **0,3** |
| Кількість вкладених у захисту коштів, $Y$ | 0,1 | 0,1 | **0,1** | 0,1 | 0,1 | **0,1** | 0,1 | 0,1 | **0,1** |
| Кількість вилученої інформації, $i$ | 0,5 | 0,333 | **0,25** | 0,667 | 0,6 | **0,571** | 0,75 | 0,714 | **0,7** |
| Ефективність захисту, $E$ | 5 | 6,667 | **7,5** | 3,333 | 4 | **4,286** | 2,5 | 2,867 | **3** |
| Цільова функція, $W_1$ | 1,04 | 1,49 | **1,715** | 0,59 | 0,77 | **0,847** | 0,365 | 0,461 | **0,5** |
| $W_2$ | 2,02 | 2,737 | **3,095** | 1,303 | 1,59 | **1,713** | 0,945 | 1,099 | **1,16** |

Одержані результати залежать від виду функції $\overline{p_i}(x,y)$. Для визначення чутливості результатів до виду цієї функції використаємо залежність $\overline{p_i}(x,y)$ у вигляді 2 (рис.3.5):

$$\overline{p_i} = \frac{\left(x_i / y_i\right)^4}{\left(x_i / y_i\right)^4 + 2^4}.$$

Результати розрахунків подано у табл.3.2.

Таблиця 3.2

Результати розрахунків показників СЗІ при використанні залежності $\overline{p_i}(x,y)$ у вигляді 2 (рис. 3.5)

| Показники | Значення показників залежно від кількості ресурсів нападу та кількості перешкод | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $X = 0{,}1$ | | | $X = 0{,}2$ | | | $X = 0{,}3$ | | |
| | Схема 1 | Схема 2 | Схема 3 | Схема 1 | Схема 2 | Схема 3 | Схема 1 | Схема 2 | Схема 3 |
| $y$ | 1 | 0,5 | 0,33 | 1 | 0,5 | 0,33 | 1 | 0,5 | 0,33 |
| $\overline{p_i}$ | $\overline{p_1}$=0,059 | $\overline{p_1}$=0,5<br>$\overline{p_2}$=0,059 | $\overline{p_1}$=0,835<br>$\overline{p_2}$=0,5<br>$\overline{p_3}$=0,059 | $\overline{p_1}$=0,5 | $\overline{p_1}$=0,94<br>$\overline{p_2}$=0,835 | $\overline{p_1}$=0,988<br>$\overline{p_2}$=0,975<br>$\overline{p_3}$=0,94 | $\overline{p_1}$=0,835 | $\overline{p_1}$=0,988<br>$\overline{p_2}$=0,975 | $\overline{p_1}$=0,998<br>$\overline{p_2}$=0,996<br>$\overline{p_3}$=0,993 |
| $R$ | 0,941 | 0,971 | **0,975** | **0,5** | 0,214 | 0,094 | **0,165** | 0,037 | 0,013 |
| $S$ | 0,1 | 0,1 | **0,1** | **0,1** | 0,1 | 0,1 | **0,1** | 0,1 | 0,1 |
| $\overline{V}$ | 0,059 | 0,029 | **0,025** | **0,5** | 0,786 | 0,906 | **0,835** | 0,963 | 0,987 |
| $E$ | 9,412 | 9,706 | **9,754** | **5** | 2,141 | 0,935 | **1,649** | 0,369 | 0,129 |
| $W_1$ | 2,231 | 2,311 | **2,324** | **1,04** | 0,268 | 0,058 | **0,135** | -0,211 | -0,275 |
| $W_2$ | 3,917 | 4,044 | **4,064** | **2,02** | 0,79 | 0,272 | **0,579** | 0,028 | -0,075 |

Як видно з таблиці 3.2, оптимальною при $X = 0{,}2$ і $X = 0{,}3$ виявляється схема 1, а не схема 3 (з розрахункових формул видно, що оптимум буде досягатись по всіх показниках одночасно, оскільки всі вони визначаються величиною $\prod_{i=1}^{m} \overline{p_i}$). Значення X, при якому оптимум переходить від схеми 3 до схеми 2, становить $X = 0{,}103$, від схеми 2 до схеми 1 – при $X = 0{,}116$.

Зазначимо, що при використанні лінійної залежності $\overline{p_i}(x,y) = \frac{1}{3} \cdot \frac{x}{y}$ (крива 3 на рис. 3.5) оптимальною стратегією захисту при $X = 0{,}1$ являється використання схем 2 та 3, при $X = 0{,}2$ - схеми 1, а при $X \geq 0{,}3$ - значення цільової функції для всіх схем становить $W_1 = -0{,}31$, $W_2 = -0{,}13$.

*Приклад 1*. Розглянемо багаторубіжну систему, до складу якої входить firewall, антивірусне програмне забезпечення, антиспам фільтр, засоби шифрування даних. Об'єктами захисту є файл-сервер та поштовий сервер. Для вилучення інформації зловмисник повинен спочатку обійти захист firewall, а потім, в залежності від об'єкту нападу, - здійснити вірусну атаку чи визначити ключ шифрування та використати спам-напад (рис. 3.6,а).

За даними [23] надійність firewall, антивірусного програмного забезпечення, антиспам фільтру та шифрування даних дорівнюють *0,484*; *0,8*; *0,9* та *0,43* відповідно. Вартість системи захисту інформації приймемо *10%* від загальних витрат компанії. Об'єм інформації, що підлягає захисту, розподілено наступним чином: на файловому сервері міститься *40%* інформації, а на поштовому - *60%*. Такий розподіл обумовлений надійністю елементів, що входять до системи захисту. При відсутності системи захисту об'єкт інформаційної діяльності вважається повністю незахищеним.

Для схеми (рис. 3.6,а) розраховані показники мають значення:

- надійність $R = 1 - \overline{p_1} \cdot (1 - \overline{q_1} \cdot (1 - \overline{q_2} \cdot \overline{q_3})) = 0.84;$
- вартість $Y = y_1 + y_2 + y_3 = 0.1;$
- кількість вилученої інформації

$$I = \overline{p_1} \cdot (\overline{q_1} \cdot g_1 + \overline{q_2} \cdot \overline{q_3} \cdot g_2)) = 0.06;$$

- ефективність $E = \frac{p_0 \cdot g - p_1 \cdot (q_1 \cdot g_1 + q_2 \cdot q_3 \cdot g_2))}{Y} = 9.40.$

*Приклад 2*. Розглянемо схему з аналогічним складом системи захисту, але з іншим розташуванням елементів (рис. 3.6,б ). Для того, щоб зловмисник вилучив інформацію з файл-серверу, йому потрібно обійти захист firewall, а потім визначити ключ

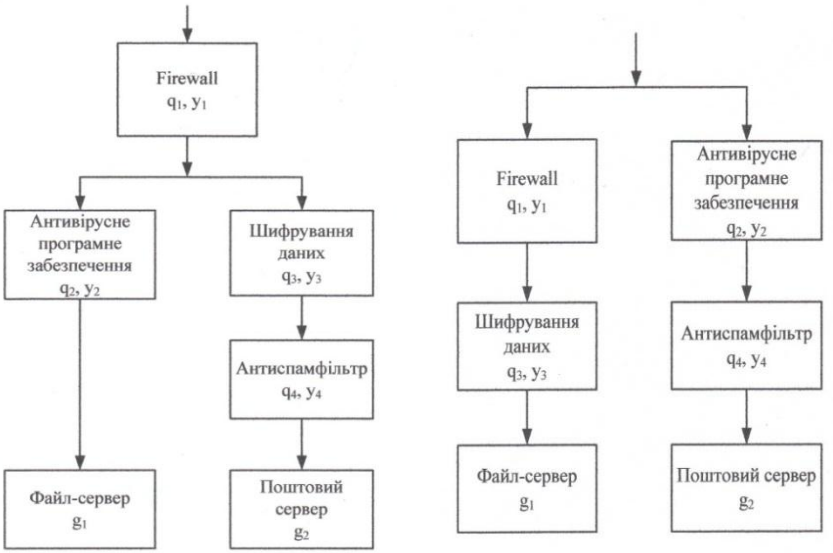
шифрування. Для отримання інформації з поштового серверу нападнику необхідно здійснити вірусну та спам-атаку.

Розраховано показники схеми (рис. 3.6,б ):

- надійність $R = (1 - \overline{q_1} \cdot \overline{q_3}) \cdot (1 - \overline{q_2} \cdot \overline{q_4})) = 0.61;$
- вартість $Y = y_1 + y_2 + y_3 = 0.1;$
- кількість вилученої інформації

$$I = \overline{q_1} \cdot \overline{q_3} \cdot g_1 + \overline{q_2} \cdot \overline{q_4} \cdot g_2 = 0.13;$$

- ефективність $E = \dfrac{\overline{p_0} \cdot g - (\overline{q_1} \cdot \overline{q_3} \cdot g_1 + \overline{q_2} \cdot \overline{q_4} \cdot g_2)}{Y} = 8.70.$




а) б)

Рис. 3.6. Схеми розташування елементів багаторубіжної системи захисту інформації

Таким чином, розглянувши різні варіанти розташування елементів багаторубіжної системи захисту інформації та розрахувавши основні показники (надійність, кількість вилученої інформації та ефективність), можна спроектувати більш ефективну та надійну систему.

## 3.2 Експертні оцінки

Як зазначалось, одним з важливих економічних завдань менеджменту інформаційної безпеки являється оптимізація розподілу ресурсів між об'єктами захисту інформації. Рішення цієї задачі викликає ряд проблем, які мають як функціональний, так і обчислювальний характер. До ключових проблем слід віднести пошук цільової функції $i(x,y)$, яка визначає кількість вилученої інформації через ресурси $x$ і $y$, спрямовані на кожний об'єкт нападом і, відповідно, захистом, кількість інформації на об'єкті і імовірнісні показники. Обчислювальні труднощі пов'язані зі значною кількістю змінних величин, частина яких (керовані змінні) задається захистом і може бути визначена з певною похибкою, а інша частина (некеровані змінні) задається нападом – її можна лише оцінити з деякою імовірністю. Розглянемо цільову функцію (4).

Величина $g_k$, яка входить як параметр в розрахунок, відноситься до керованих змінних, але її визначення викликає певні труднощі через недостатню розробку методики розрахунку кількості інформації. Величини $p_k$ і $q_k(x,y)$ – некеровані змінні, і їх визначення ще більш проблематичне.

За браком статистичної інформації для визначення змінних, які вводяться в розрахунок, і вибору залежності $f_k(x,y)$, яка в найбільшій мірі відповідає реальним ситуаціям, доводиться звертатись до методу експертних оцінок.

Проілюструємо його на прикладі системи захисту інформації, яка складається з $l$об'єктів і оцінюється групою з $m$ експертів.

На першому етапі оцінимо значення параметрів $g_k$ на кожному з об'єктів. В табл. 4.1 приведені значення величин, одержаних в результаті експертних оцінок:

$G_{nk}$ – нормовані до 1 вагові коефіцієнти $k$–го об'єкта, визначені $n$–им експертом;

$\overline{G_k}$ – усереднене по всім експертам значення коефіцієнта $G_{nk}$,

$$\overline{G_k} = \frac{\sum_{n=1}^{m} G_{nk}}{m}, \quad m = 5;$$

$\Delta G_{nk} = G_{nk} - \overline{G_k}$ – відхилення $G_{nk}$ від середнього значення;

$\sum_{k=1}^{l} |\Delta G_{nk}|$ – сума абсолютних значень відхилень оцінок кожного експерта від середніх значень по всім об'єктам;

$c_n$ – ваговий коефіцієнт n-го експерта, який визначається, як

$$c_n = \frac{1}{\sum_{k=1}^{l} |\Delta G_{nk}|};$$

$\sigma_k^2 = \sum_{n=1}^{m} \frac{\Delta G_{nk}^2}{l-1}$ – дисперсія оцінок параметра $G_{nk}$, даного експертами для $k$-го об'єкта;

$V_k = \frac{\sigma_k}{\overline{G_k}}$ – варіація оцінок $G_{nk}$ для $k$–го об'єкта.

Величину $\sum_{k=1}^{3} \Delta G_{nk}$ можна розглядати як показник кваліфікації $n$-го експерта, а $\Delta G_{nk}$ – як ступінь обізнаності$n$-го експерта з $k$-тим об'єктом. Після нормування встановлюємо остаточно показники кваліфікації експертів, позначаючи їх $c_n$.

В табл. 4.2 приведені значення $c_n\, G_{nk}$ – параметра важливості $k$-го об'єкта, дані $n$–им експертом, з врахуванням його кваліфікації, а також остаточні значення параметра $g_k$, який характеризує об'єм інформації на $k$–му об'єкті (тобто його важливість) і визначається як адитивна функція показників $G_{nk}$ з ваговими коефіцієнтами $c_n$:

$$g_k = \sum_{n=1}^{m} c_n G_{nk} \qquad (2.5)$$

Табл.3.3

Результати експертних оцінок

| Номер об'єкта $k$ / показник / Номер експерта $n$ | 1 | | 2 | | 3 | | $\sum_{k=1}^{3} \lvert\Delta G_{nk}\rvert$ | $c_n$ |
|---|---|---|---|---|---|---|---|---|
| | $G_{n1}$ | $\Delta G_{n1}$ | $G_{n2}$ | $\Delta G_{n2}$ | $G_{n3}$ | $\Delta G_{n3}$ | | |
| 1 | 0,47 | -0,07 | 0,33 | 0,05 | 0,20 | 0,02 | 0,14 | 0,18 |
| 2 | 0,50 | -0,04 | 0,25 | -0,03 | 0,25 | 0,07 | 0,14 | 0,18 |
| 3 | 0,50 | -0,04 | 0,28 | 0,00 | 0,22 | 0,04 | 0,08 | 0,32 |
| 4 | 0,71 | 0,17 | 0,21 | -0,07 | 0,07 | -0,11 | 0,35 | 0,07 |
| 5 | 0,50 | -0,04 | 0,33 | 0,05 | 0,17 | -0,01 | 0,10 | 0,25 |
| $\bar{G}_k$ | 0,54 | | 0,28 | | 0,18 | | | |
| $\sigma_k$ | 0,098 | | 0,051 | | 0,069 | | | |
| $V_k$ | 0,181 | | 0,185 | | 0,380 | | | |

Табл.3.4

Розрахунок показників $g_k$

| $c_n G_{nk}$ / $n$ | $c_n G_{n1}$ | $c_n G_{n2}$ | $c_n G_{n3}$ |
|---|---|---|---|
| 1 | 0,085 | 0,059 | 0,036 |
| 2 | 0,090 | 0,045 | 0,045 |
| 3 | 0,160 | 0,093 | 0,070 |
| 4 | 0,050 | 0,015 | 0,005 |
| 5 | 0,125 | 0,083 | 0,042 |
| $g_k$ | 0,50 | 0,30 | 0,20 |

Цим остаточно встановлено визначений експертами розподіл інформації по об'єктах: $\{g_k\}=0,5:0,3:0,2$. Аналогічно розраховуються параметри $p_k$ в (4).

Після визначення параметрів цільової функції звертаємось до вибору залежностей $g(x,y)$і$f(x,y)$. Припустимо, що кожен з п'яти експертів запропонував для опису певного об'єкта такі функції$f^{(n)}(x)$(для спрощення покладемо$y=1$):

1) $f_k^{(1)}(x)=\dfrac{x}{x+1}$

2) $f_k^{(2)}(x)=\dfrac{x^2}{x^2+2}$

3) $f_k^{(3)}(x)=\dfrac{x^4}{x^4+4}$

4) $f_k^{(4)}(x) = 1 - e^{-x^2}$

5) $f_k^{(5)}(x) = 1 - e^{-x^4}$

Результуючу залежність *f(x,y)* побудуємо як адитивну функцію:

$$f_k(x) = c_1 f_k^{(1)} + c_2 f_k^{(2)} + c_3 f_k^{(3)} + c_4 f_k^{(4)} + c_5 f_k^{(5)} \qquad (2.6)$$

Ця залежність зображена на рис. 3.7 неперервною лінією (крива 6).

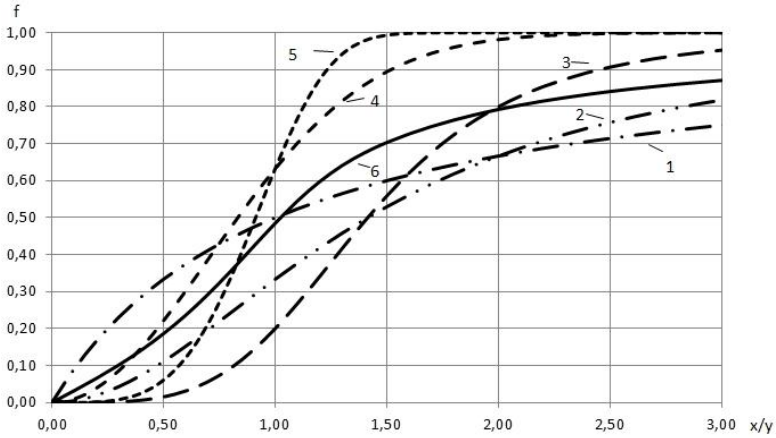



Рис. 3.7. Залежність втрат інформації від співвідношення ресурсів

Знайшовши за допомогою (2.5), (2.6) значення параметрів $g_k$ і вид функцій $f_k(x)$, можемо розрахувати значення цільової функції (2.1).

## Контрольні питання

1. Відмінність розрахунку показників в багатоелементних радіоелектронних системах і в багаторубіжних системах захисту інформації.
2. Схеми з'єднання елементів в багаторубіжних системах і її показники.
3. Загальний вигляд схеми багаторубіжної системи і розрахункові формули.
4. Вигляд цільової функції при різних значеннях вагових коефіцієнтів і розрахунок показників.

5. Вибір оптимального розташування елементів багаторубіжної системи.

6. Сутність метода експертних оцінок. Показники кваліфікації експертів і ступеня їх обізнаності.

7. Схема розрахунку кількості інформації на об'єктах за оцінками експертів.

8. Розрахунок залежності $f(x)$ на основі оцінок експертів.

## IV. ЗАСТОСУВАННЯ МАТЕМАТИЧНИХ МЕТОДІВ В ЕКОНОМІЧНИХ ЗАДАЧАХ

Розглянемо спочатку економічні задачі загального змісту[2,3], а потім спробуємо застосувати приведені методики до задач інформаційної безпеки.

### 4.1 Диференціальне і інтегральне числення

#### *4.1.1 Оптимізація випуску продукції*

**А.** Виробник реалізує за один день відеокамери кількістю $q$ по ціні $p$ за одиницю, а витрати задаються залежністю:

$$S(q) = aq+\lambda q^2, \qquad де\ 0<\lambda<1, \qquad a<p.$$

Знайти оптимальний обсяг випуску продукції, відповідний прибуток і рентабельність виробництва.

Валовий прибуток від реалізації продукції – це різниця між валовим доходом $pq$ від реалізації продукції і витратами $S(q)$ на виробництво:

$$D(q) = pq – (aq +\lambda q^2).$$

Для знаходження оптимуму використовуємо теорему Ферма:

$$D'(q)=(p – a) – 2\lambda q_0 =0,$$

звідки оптимальний обсяг випуску продукції становить:

$$q_0 = (p – a)/2\lambda.$$

Друга похідна $D''(q) = - 2\lambda < 0$, отже $q_0$ – точка максимуму.

Максимальний розмір прибутку:

$$D(q_0) = (p - a)^2/2\lambda – (\lambda(p - a)^2)/4\lambda^2 = (p – a)^2/4\lambda.$$

Рентабельність виробництва:

$$R = D/S= (p - a)^2/4\lambda(aq_0+ \lambda q_0^2).$$

Розглянемо числовий приклад, в якому величини $p$, $a$, $\lambda$, $D$виражаються в грошових одиницях (*г.о.*).

Дано: $p = 10\,000$

$a = 6\,000$

$\lambda = 100$

$q_0 = (p - a)/2\lambda = 4\,000/200 = 20$

$D(q_0) = (p - a)^2/4\lambda = (4 \cdot 10^3)^2/400 = 40\,000$

$R = (p - a)^2/4\lambda(aq_0 + \lambda q_0^2) = (4 \cdot 10^3)^2/400(6 \cdot 20 \cdot 10^3 + 100 \cdot 400 = 0.25$, *або 25%.*

**В.** Фірма планує випускати генератори шуму. Вивчення ринку показало, що залежність попиту $q$ від ціни $p$ задається рівністю:

$$q = 100\,000 - 200p,$$

де $q$ – кількість генераторів для продажу за рік.

Витрати фірми в грошових одиницях на випуск $q$ генераторів становлять:

$$S(q) = 500 + 100q + 0{,}003q^2.$$

Розрахувати прибуток від продажу генераторів і визначити його оптимальне значення.

Валовий дохід: $R = p \cdot q$.

З першої рівності виразимо $p$ через $q$: $p = 500 - 0{,}005q$.

Отже, валовий дохід $R$ залежить від випуску $q$ продукції так:

$$R(q) = q(500 - 0{,}005q).$$

Вираховуючи з валового доходу затрати на виробництво, одержимо валовий прибуток:

$D(q) = R - S = 500q - 0.005q^2 - 15000 - 100q - 0.003q^2 = -0{,}008q^2 + +400q - 15000.$

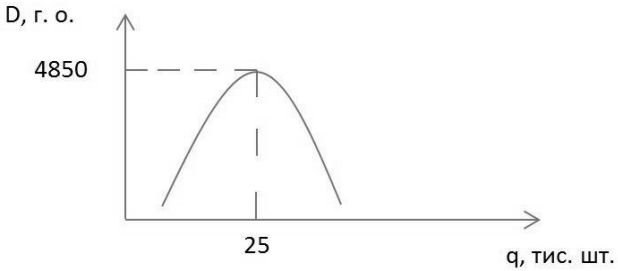



Рис.4.1 Залежність прибутку від обсягу виробництва

Знайдемо точку екстремуму:

$$D'(q) = 400 - 0.016q = 0,$$

звідки оптимальний випуск продукції становить (рис. 6.1):

$$q_0 = \frac{400}{0.016} = 25000,$$

а оптимальний прибуток:

$$D(q_0) = 4850000.$$

### *4.1.2 Крива Лоренца*

Залежність частини $d$ доходів від частини $k$ населення, що їх одержує (крива Лоренца[1]) характеризує ступіть нерівності в розподілі доходів населення(рис. 6.2).

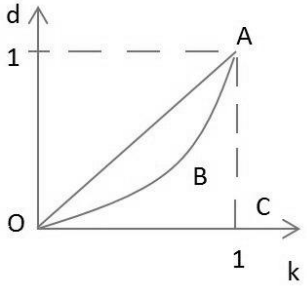


Рис. 4.2 Крива Лоренца

При рівномірному розподілі крива Лоренца вироджується в пряму. Тому ступінь нерівності в доходах кількісно характеризують коефіцієнтом Джині[2] *D*– це відношення площі фігури *OAB* до площі трикутника *OAC*.

Приведемо чисельний приклад.

За даними соціологічних досліджень 90% населення однієї з країн одержує приблизно60% національного доходу, а інші 10%–40% доходу.

За таких даних крива Лоренца може бути апроксимована функцією $d = 1 - \sqrt{1 - k^2}$. Перевіримо це твердження. Підставимо $k = 0.9$ у вираз функції значення. Одержимо $d = 1 - \sqrt{1 - 0.89} = 0.56$, що близько до значення,приведеного в умові (приблизно 60%).

---

[1]Макс Лоренц (1876 – 1959) – американський економіст
[2]Коррадо Джині (1884 – 1965) – італійський економіст

Розрахуємо коефіцієнт Джині $D = \frac{S_{OAB}}{S_{OAC}}$. Очевидно, що $S_{OAC} = \frac{1}{2}$. Площу фігури *OAB* знайдемо як різницю $S_{OAB} = S_{OAC} - S_{OBAC}$. Площа фігури *OBAC* знаходиться інтегруванням:

$$S_{OBAC} = \int_0^1 (1 - \sqrt{1-k^2})dk = \int_0^1 dk - \int_0^1 \sqrt{1-k^2}\,dk.$$

$$\int_0^1 \sqrt{1-\mathrm{k}^2}\,\mathrm{dk} = \left|\begin{matrix}\mathrm{k} = \sin t \\ \mathrm{dk} = \cos t\,\mathrm{dt}\end{matrix}\right| = \int_0^{\pi/2} \sqrt{1-\sin^2 t}\cdot\cos t\,\mathrm{dt} =$$

$$= \int_0^{\pi/2} \cos^2 t\,\mathrm{dt} = \int_0^{\pi/2} \frac{1-\cos 2t}{2}\,\mathrm{dt} = \frac{\pi}{4} + \frac{1}{4}\sin 2t\Big|_0^{\pi/2} = \frac{\pi}{4}$$

$$\mathrm{S_{OBAC}} = 1 - \frac{\pi}{4}$$

$$\mathrm{S_{OAB}} = \frac{1}{2} - 1 + \frac{\pi}{4} = \frac{\pi}{4} - \frac{1}{2}$$

Звідси коефіцієнт Джині:

$$D = \frac{S_{OAB}}{S_{OAC}} = \frac{\frac{\pi}{4} - \frac{1}{2}}{\frac{1}{2}} = \frac{\pi}{2} - 1 = 0.57.$$

Коефіцієнт має досить високе значення.

## 4.2 Диференціальні рівняння

### *4.2.1 Ефективність реклами*

Припустимо, що ми реалізуємо продукцію (виробничу, фінансову, інформаційну), про яку на поточний момент з числа потенційних покупців $N$ знає лише $x$ покупців. Реклама про продукцію розповсюджується через ЗМІ, а надалі – в результаті спілкування покупців між собою. Вважаємо, що швидкість зростання кількості обізнаних про продукцію пропорційна кількості $x$ знаючих про неї, і кількості $N$-$x$ покупців, які ще не знають про неї:

$$\frac{dx}{dt} = kx(N - x),$$

де $k$ – коефіцієнт пропорційності.

Початкові умови: $t = 0$ $x_0 = \frac{N}{\gamma}$, де $\gamma > 1$.

Відокремлюємо змінні в диференціальному рівнянні:

$$\frac{dx}{x(N - x)} = kdt.$$

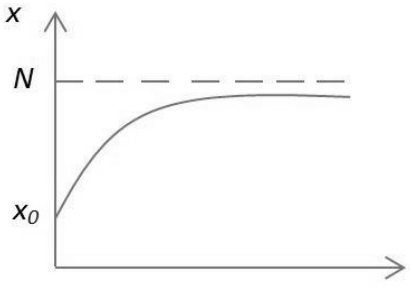
Використовуючи метод неозначених коефіцієнтів, зводимо нелінійний дріб до суми лінійних:

$$\frac{1}{x(N - x)} = \frac{A}{x} + \frac{B}{N - x} = \frac{A(N - x) + Bx}{x(N - x)}$$

$$\begin{cases} -A + B = 0 \\ AN = 1 \end{cases} \Bigg| \begin{matrix} B = A = \frac{1}{N} \\ A = \frac{1}{N} \end{matrix}$$

Підставляємо коефіцієнти $A$, $B$ і виконуємо інтегрування:

$$\int \frac{dx}{x(N - x)} = \frac{1}{N}\left[\int \frac{dx}{x} + \int \frac{dx}{N - x}\right.$$



Після інтегрування маємо:

$$\frac{1}{N} ln \frac{x}{N-x} = kt + C$$

$$\frac{x}{N-x} = e^{Nkt} \cdot e^{NC} = Ae^{Nkt},$$

де $A = e^{NC}$.

Розв'язуємо останнє рівняння відносно $x$ і отримуємо:

$$x = N \frac{Ae^{Nkt}}{Ae^{Nkt} + 1} = \frac{N}{1 + Pe^{-Nkt}},$$

де $P = \frac{1}{A}$.

Рис. 4.3 Залежність кількості обізнаних покупців від часу

Отже, $x = \frac{N}{1 + Pe^{-Nkt}}$ - це рівняння логістичної кривої(рис.4.3).

Враховуємо початкові умови і визначимо невідомі константи:

$$t = 0 \; \frac{N}{\gamma} = \frac{N}{1+P},$$

звідки $P = \gamma - 1$.

Остаточно: $x = \frac{N}{1 + (\gamma - 1)e^{-Nkt}}$.

Коефіцієнт $k$ визначається емпірично.

### *4.2.2 Зростання обсягу продукції з часом*

**А.** В умовах ненасиченого ринку.

Нехай *q(t)*- кількість продукції,*p* - ціна продукції, випущеної за час *t*.

Позначимо через $I(t)$ суму інвестицій, тобто частку доходу, витрачену на розширення виробництва:

$$I(t) = mpq(t), \qquad (4.1)$$

де $m$ -норма інвестицій ( $0 < m < 1$).

Припустимо, що ринок ненасичений, тобто весь товар, що поступає на ринок, реалізується. Надходження інвестицій призведе

(4.2)

до збільшення випуску продукції, причому це збільшення пропорційне кількості інвестицій:

$$q' = \eta I ,$$

де $q' = \frac{\partial q}{\partial t}$, $\eta$ - коефіцієнт пропорційності.

Підставимо (4.1) в (4.2) і одержимо:

$$q' = kq, \qquad (4.3)$$

де $k = \eta mp$.

Це лінійне диференціальне рівняння з відокремлюваними змінними. Проінтегруємо його:

$$\frac{\partial q}{q} = k\partial t$$

$$ln|q| = kt + ln\,C,\ q = Ce^{kt}.$$

Стала $C$ визначається з умови

$q(0) = q_0$, звідки $q_0 = C$.

Остаточно: $q = q_0 e^{kt}$. Цей вираз відображає необмежене зростання кількості продукції з часом (рис. 4.4). Звичайно ця залежність буде діяти тільки в межах ненасиченого ринку.

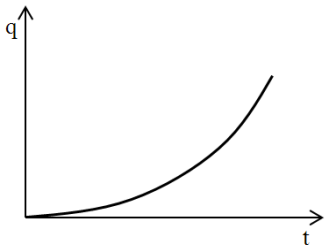



Рис. 4.4 Залежність кількості реалізованої продукції від часу в умовах ненасичено го ринку

**В.** В умовах насиченого ринку.

Враховуємо тепер, що із збільшенням випуску продукції буде відбуватися насичення ринку. При цьому ціна вже не буде сталою величиною, а буде зменшуватись, при чому буде виражатись спадною функцією $p(q)$:

$$\partial p \big/ \partial q < 0 .$$

Тоді диференціальне рівняння *(3)* приймає вид:

$$q' = \gamma p(q)q,$$

де $\gamma = \eta m$.

Припустимо, що залежність $p(q)$ має лінійний характер:

$$p(q) = b - aq.$$

Тоді $q' = \gamma(b - aq)q.$

Це лінійне неоднорідне диференціальне рівняння першого порядку зі сталими коефіцієнтами.

Відокремлюємо змінні:

$$\frac{dq}{q(b-aq)} = kdt; \qquad \int \frac{dq}{q(b-aq)} = kt + C_1.$$

Застосовуємо метод неозначених коефіцієнтів і переходимо від дробно-нелінійного дробу відносно $q$до суми дробно-лінійних дробів:

$$\frac{1}{q(b-aq)} = \frac{A}{q} + \frac{B}{b-aq} = \frac{A(b-aq)+Bq}{q(b-aq)}.$$

Прирівнюємо чисельники:

$$A(b-aq)+Bq = 1.$$

Прирівнюємо коефіцієнти при однакових степенях $q$:

$$\begin{cases} -Aa + B = 0 \\ Ab = 1 \end{cases} \quad \begin{vmatrix} B = \frac{a}{b} \\ A = \frac{1}{b} \end{vmatrix}$$

Підставляємо знайдені значення $A$ і $B$ та інтегруємо:

$$\int \frac{dq}{q(b-aq)} = \frac{1}{b}\int\left(\frac{1}{q} + \frac{a}{b-aq}\right)dq = \frac{1}{b}\left[\ln q - \ln(b-aq)\right] =$$

$$= \frac{1}{b}\ln\frac{q}{b-aq},$$

$$\frac{1}{b}\ln\frac{q}{b-aq} = kt + \ln C,$$

$$\frac{q}{b-aq} = C \cdot e^{bkt},$$

$$q = \frac{Cbe^{bkt}}{1+Cae^{bkt}}$$

і остаточно:

$$q = \frac{Cb}{e^{-bkt} + Ca}.$$

Ця залежність зображується логістичною кривою (рис.4.5).

При збільшенні інвестицій кількість реалізованої продукції буде зростати до рівноважного значення $q=b/a$, яке визначається насиченням ринку.

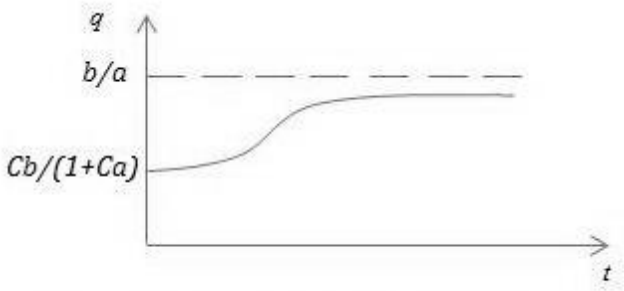



Рис. 4.5 Залежність кількості реалізованої продукції від часу в умовах насиченого ринку.

### *4.2.3 Встановлення рівноважної ціни*

Зміна ціни з часом при її прагненні до рівноважного значення може приймати різні форми, зокрема вона може мати як релаксаційний, так і коливальний характер. Розглянемо ці варіанти.

**А.** Нехай попит і пропозиція змінюються з часом відповідно до наступних залежностей:

функція попиту $y(t) = 4p' - 2p + 7,$

функція пропозиції $z(t) = 8p' + 2p - 5,$

де $p(t)$ – ціна продукції.

Початкові умови: t = 0; $p(0)$ =6.

Рівноважна ціна встановлюється, коли попит і пропозиція стануть рівними:

$$y(t) = z(t).$$

Тоді:

$$8p' + 2p - 5 = 4p' - 2p + 7.$$

Одержуємо диференціальне рівняння відносно $p(t)$:

$4p' + 4p - 12 = 0$, звідки:

$$p' + p - 3 = 0,\ p' = -p + 3,\ \frac{dp}{dt} = -p + 3.$$

Відокремлюємо змінні:

$$\frac{dp}{p-3} = -dt,\ \int \frac{dp}{p-3} = -\int dt + \ln C,\quad ln(\ p-3) = -t + \ln C,$$

$$l\,n(\frac{p-3}{C}) = -t.$$

$p(y) = Ce^{-t} + 3.$ При $t = 0$ і $p(0)=6$, звідки $C = 3$, і остаточно:

$p(t) = 3e^{-t} + 3.$

Ціна з часом буде змінюватись по закону:

$$p(t) = 3e^{-t} + 3,$$

Прагнучи до рівноважного значення $p_0 = 3$ (рис.4.6).

Ця залежність має релаксаційний характер, який визначається видом функцій попиту і пропозиції.

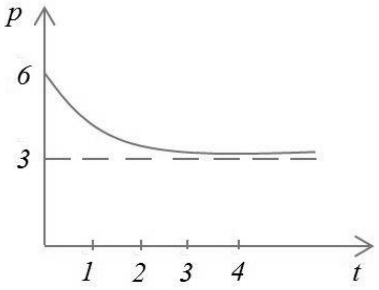



Рис. 4.6 Релаксаційний характер залежності ціни від часу

**В**. Зміна попиту і пропозиції з часом описуються залежностями:
функція попиту $y(t) = 3p'' - p' - 2p + 18,$
функція пропозиції $z(t) = 4p'' + p' + 3p + 3.$
Початкові умови: $t = 0$ => $p(0) = 5;\ p'(0) = 2.$
Використовуємо умову рівноваги попиту і пропозиції:

$$y(t) = z(t).$$

Тоді:

$$3p'' - p' - 2p + 18 = 4p'' + p' + 3p + 3$$

$$p'' + 2p' + 5p = 15$$

Отримане рівняння − лінійне неоднорідне диференціальне рівняння другого порядку зі сталими коефіцієнтами. Його загальний розв'язок − це сума загального розв'язку відповідного однорідного рівняння і частинного розв'язку неоднорідного рівняння.

а) Знайдемо розв'язок однорідного диференціального рівняння $p'' + 2p' + 5p = 0.$

Характеристичне рівняння: $\lambda^2 + 2\lambda + 5 = 0.$

Корені рівняння:

$$\lambda_{12} = \frac{-1 \pm \sqrt{4-20}}{2} = -1 \pm 2i \quad \begin{bmatrix} \lambda_1 = -1 + 2i \\ \lambda_2 = -1 - 2i \end{bmatrix}$$

Загальний розв'язок:

$$\mathrm{p} = \mathrm{C_1 e}^{\lambda_1 \mathrm{t}} + \mathrm{C_2 e}^{\lambda_2 \mathrm{t}} = \mathrm{C_1 e}^{(-1+2\mathrm{i})\mathrm{t}} + \mathrm{C_2 e}^{(-1-2\mathrm{i})\mathrm{t}} = \mathrm{C_1 e}^{-\mathrm{t}}\mathrm{e}^{2\mathrm{it}} +$$

$$+ \mathrm{C_2 e}^{-\mathrm{t}}\mathrm{e}^{-2\mathrm{it}} = \mathrm{e}^{-\mathrm{t}}(\mathrm{C_1 cos2t} + \mathrm{C_1 isin2t} + \mathrm{C_2 cos2t} - \mathrm{C_2 isin2t}) =$$

$$= \mathrm{e}^{-\mathrm{t}}(\mathrm{Acos2t} + \mathrm{Bsin2t}) = \mathrm{Ae}^{-\mathrm{t}}\mathrm{cos2t} + \mathrm{Be}^{-\mathrm{t}}\mathrm{sin2t},$$

де $A = C_1 + C_2$, $B = i(C_1 - C_2)$.

б) Шукаємо частинний розв'язок неоднорідного рівняння $p'' + 2p' + 5p = 15.$

Неважко збагнути що розв'язком може бути $p = C,$ де $C$ - невідома стала.

Знайдемо її, підставляючи розв'язок в рівняння:

$5C = 15, C = 3.$

Отже,загальний розв'язок неоднорідного диференціального рівняння має вигляд:

$$p = Ae^{-t} \cos 2t + Be^{-t} \sin 2t + 3.$$

Невідомі сталі $A$ і $B$знаходимо з початкових умов:

$$t = 0, \qquad p(0) = 5,\ p'(0) = 2.$$

$$p(0) = A + 3 = 5,\ A = 2.$$

$$p'(0) = \left[ -Ae^{-t} \cos 2t - 2Ae^{-t} \sin 2t - Be^{-t} \sin 2t + Be^{-t} \cos 2t \right] =$$

$$= -A + D = 2$$

$$B = 4$$

Підставляємо одержані значення $A$ і $B$ у вираз для загального розв'язку і проаналізуємо зміну ціни з часом:

$$p(t) = 2e^{-t}cos2t + 4e^{-t}sin2t + 3.$$

$t = 0$ $\quad p = 5$

$t = \frac{\pi}{2}$ $\quad p = -2e^{-\pi/2} + 3 \cong 2.8$

$t = \pi$ $\quad p = 2e^{-\pi} + 3 \cong 3.09$

$t = \frac{3\pi}{2}$ $\quad p = -4e^{-3\pi/2} + 3 \cong 2.96$

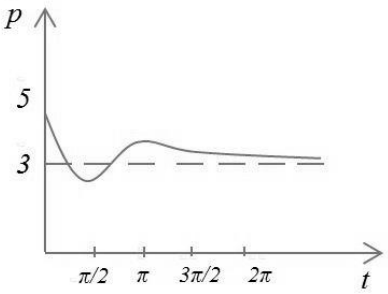



Рис. 4.7 Коливальний характер залежності ціни від часу

$t = 2\pi \quad p = 2e^{-2\pi} + 3 = 3$

$t \to \infty \quad p \to 3$

Залежність $p(t)$ має коливальний характер (рис.4.7). Це обумовлено появою в функціях попиту і пропозиції другої похідної.

## Контрольні питання

1. Критерій наявності оптимального значення, його застосування.
2. Крива Лоренца, її сутність. Методика розрахунку коефіцієнта Джині.
3. Приклади застосування диференціальних рівнянь до розгляду економічних задач.

# V. ЗАСТОСУВАННЯ СТАТИСТИЧНИХ МЕТОДІВ В ЗАДАЧАХ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

## 5.1 Основні поняття і показники математичної статистики

Математична статистика – це розділ математики, присвячений вивченню закономірностей, які мають місце в масових явищах і статистичних сукупностях (множинах).

Статистичною сукупністю називається сукупність об'єктів, які на рівні зі схожими мають відмінні (варіюючи) ознаки, за якими і проводиться статистичне дослідження. Варіюючи ознаки поділяються на атрибутивні (якісні) і кількісні. Атрибутивні ознаки не піддаються числовому виразу. Значення, які приймають кількісні ознаки, називаються варіантами. За характером варіювання кількісні ознаки поділяються на дискретні і неперервні. Залежно від повноти обстеження розрізняють генеральну (загальну) і вибіркову (часткову) сукупності. Вибіркова сукупність вибирається з генеральної для спостереження. Статистичні сукупності, розчленовані за однією ознакою, називають одновимірними, за двома - двовимірними, за трьома і більше - багатовимірними. Сукупності можна поділити на окремі групи за групувальними ознаками. Число одиниць сукупності, які мають однакове значення ознаки, називають частотою даної варіанти, або її вагою. Систематизація статистичних даних може вестись шляхом її розміщення у зростаючому або спадаючому порядку. При цьому одержуємо ранжируваний ряд розподілу. Ранжируваний ряд із зазначенням для кожного інтервалу (групи) відповідних частот називається варіаційним рядом.

Комплексний опис статистичних розподілів включає в себе показники:

- центру розподілу;
- ступеню варіації;
- форми розподілу.

Центр розподілу характеризують різного роду середні величини, мода, медіана. Серед середніх величин найбільше розповсюдження мають арифметична, квадратична, геометрична.

За способом обчислення всі середні ділять на прості і зважені. Розрахункові формули для середніх величин приведені в таблиці.

| № | Середня величина | Проста | Зважена |
|---|---|---|---|
| 1 | Арифметична | $\bar{x} = \frac{\sum_{i=n}^{n} x_i}{n}$ | $\bar{x} = \frac{\sum_{i=n}^{n} x_i n_i}{n}$ |
| 2 | Квадратична | $\bar{x} = \sqrt{\frac{\sum_{i=n}^{n} x_i^2}{n}}$ | $\bar{x} = \sqrt{\frac{\sum_{i=n}^{n} x_i^2 n_i}{n}}$ |
| 3 | Геометрична | $\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$ | $\bar{x} = \sqrt[n]{n_1 x_1 \cdot n_2 x_2 \cdot \ldots \cdot n_n x_n}$ |

В таблиці: $x_i$ – випадкова величина, $n_i$ – ваговий коефіцієнт, $n$ – кількість вибірок.

Для точності варіаційні ряди зображають графічно у вигляді гістограм, полігонів або кумулят. Гістограми застосовують для зображення інтервальних розподілів вибірки.

Порівняння середніх для одного і того ж варіаційного ряду приводить до їх розміщення, яке визначає мажоритарність середніх: $\bar{x}_{геом} < \bar{x}_{ад} < \bar{x}_{кв}$.

Додаткову інформацію про розподіл дають так звані структурні середні - мода і медіана. Модою *Mo* називається значення ознаки, яке має найбільшу частоту в варіаційному ряду.

Медіаною називають значення ознаки, котра поділяє ранжируваний ряд розподілу на дві рівні частини. Якщо в дискретному ряду $2m+1$ випадків, то медіаною буде значення ознаки для випадку $m+1$, тобто $x_{m+1}$. якщо в ряду маємо парне число $2m$ випадків, то $Me = \frac{x_m + x_{m+1}}{2}$.

В одномодальних симетричних рядах розподілу середня арифметична, мода і медіана співпадають: $\bar{x} = Mo = Me$.

Ще однією характеристикою структури варіаційного ряду є квартилі, які ділять ранжируваний ряд на 4 рівні частини.

Введені середні величини $\bar{x}$, *Mo*, *Me* відображають центр варіаційного ряду з різних точок зору. Кожна з них дається одним

числом, яке звичайно, не може характеризувати форму кривої розподілу, характер зміни (варіацію) і не завжди співпадає з її максимальним значенням. Основними показниками, що характеризують варіацію певної ознаки є розмах варіації, середнє лінійне відхилення, дисперсія, середнє квадратичне відхилення, квартильне відхилення, коефіцієнт варіації.

Розмах варіації є різниця між максимальним і мінімальним значеннями ознаки: $R = x_{\max} - x_{\min}$.

Розмах варіації залежить від двох крайніх значень ознаки і не враховує проміжних значень частот. Цей недолік усувається введенням двох наступних показників.

Середнє лінійне відхилення - це середнє арифметичне з абсолютних значень відхилень окремих варіант від середньої арифметичної: просте $\bar{l} = \dfrac{\sum_{i=1}^{n} \left| x_i - \bar{x} \right|}{n}$, зважене $\bar{l} = \dfrac{\sum_{i=1}^{n} \left| x_i - \bar{x} \right| \cdot n_i}{n}$.

Дисперсією $D$ називається середній квадрат відхилень усіх значень ознаки від його середньої величини:проста $D = \sigma^2 = \dfrac{\sum_{i=1}^{n} \left( x_i - \bar{x} \right)^2}{n}$, зважена $D = \sigma^2 = \dfrac{\sum_{i=1}^{n} \left( x_i - \bar{x} \right)^2 \cdot n_i}{n}$.

Середнє квадратичне відхилення $\sigma$ (його ще називають стандартним) – це корінь квадратний з дисперсії: просте $\sigma = \sqrt{\dfrac{\sum_{i=1}^{n} \left( x_i - \bar{x} \right)^2}{n}}$, зважене $\sigma = \sqrt{\dfrac{\sum_{i=1}^{n} \left( x_i - \bar{x} \right)^2 \cdot n_i}{n}}$.

При розрахунку дисперсії можна скористатись формулою:

$$\sigma^2 = \overline{x^2} - \left( \bar{x} \right)^2 = \frac{\sum_{i=1}^{n} x_i^2 n_i}{n}$$

Між $\sigma$ і $R$ в нормально розподіленій сукупності існує співвідношення: $\sigma \cong 1{,}25\bar{l} \cong \dfrac{1}{6} R$.

І, нарешті, коефіцієнт варіації – це відношення середнього квадратичного відхилення до середнього арифметичного значення: $V = \frac{\sigma}{\overline{x}}$.

Підводячи підсумок, перелічимо згадані показники.

Показники центра розподілу:

- середня арифметична;
- мода;
- медіана.

Показники варіації ознаки:

- розмах варіації;
- середнє лінійне відхилення;
- середнє квадратичне відхилення;
- дисперсія;
- коефіцієнт варіації.

Проте вказані показники не дають інформації про форму розподілу, а саме про її симетрію чи ступінь асиметрії (тут можлива лівостороння і правостороння скісність), ступінь гостровершинності (ексцесу) тощо. Для того, щоб розглянути ці показники, введемо поняття моменту розподілу.

Моментом розподілу називається середня арифметична величина з піднесених до заданої степені відхилень окремих варіант від деякої постійної величини:

$$M_k = \frac{\sum_{i=1}^{n} (x_i - A)^k \cdot n_i}{\sum_{i=1}^{n} n_i} = (x_i - A)^k,$$

де А - постійна величина, від якої визначається відхилення. Це може бути 0, середнє арифметичне або умовний початок відліку; *k* - показник степеня, який визначає порядок моменту.

Залежно від постійної величини, від якої визначається відхилення, розрізняють три види моментів: початкові $M_k$ *(A=0),* центральні $\mu_k$ *(A* $= \overline{x}$*)* та умовні $m_k(A=x_0)$. Якщо при обчисленні за

вихідну величину береться нуль, то моменти розподілу називаються початковими:

$$M_k = \overline{(x_i - 0)^k} = \frac{\sum_{i=1}^{n} x_i^k n_i}{\sum_{i=1}^{n} n_i}.$$

Якщо за вихідну величину береться відхилення від середнього арифметичного значення, то моменти називаються центральними:

$$m_k = \overline{\left(x_i - \bar{x}\right)^k} = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^k \cdot n_i}{\sum_{i=1}^{n} n_i}.$$

Якщо за вихідну величину беруться відхилення від довільно взятої величини $x_0$, тобто від умовного початку відліку, то моменти називаються умовними:

$$m_k = \overline{(x_i - x_0)^k} = \frac{\sum_{i=1}^{n} (x_i - x_0)^k \cdot n_i}{\sum_{i=1}^{n} n_i}.$$

З приведених виразів видно, що початковий момент першого порядку $M_1 = \frac{\sum x_i n_i}{\sum n_i}$ є середнє арифметичне $\bar{x}$ значень $x$і являється показником центра розподілу.

Центральний момент першого порядку завжди дорівнює нулю: $\sum \left(x_i - \bar{x}\right) = 0$. Центральний момент другого порядку $\mu_2 = \frac{\sum \left(x_i - \bar{x}\right)^2 \cdot n_i}{\sum n_i}$ визначає дисперсію: $\mu_2 = \sigma^2$. Центральний момент третього порядку $\mu_3$ дорівнює нулю в симетричному розподілі, а в несиметричному – визначає ступінь асиметрії

(скісності): $\mu_3 = \frac{\sum \left(x_i - \bar{x}\right)^3 \cdot n_i}{\sum n_i}$. Центральний момент четвертого порядку застосовується при обчисленні показника гостровершинності (ексцесу): $\mu_4 = \frac{\sum \left(x_i - \bar{x}\right)^4 \cdot n_i}{\sum n_i}$.

Оскільки моменти залежать від прийнятої системи одиниць, на практиці розглядають не абсолютне значення моментів, а їх відношення до стандартного відхилення у відповідній степені. Так, асиметрію (скісність) розподілу характеризують коефіцієнтом асиметрії (скісності), який позначається через *As* або через $k_{ck}$ і дорівнює $As = \frac{\mu_3}{\sigma^3}$. Відношення моменту *k-ого* порядку до середнього квадратичного відхилення в *k-ій* степені називається нормованим моментом. Таким чином, $As = \frac{\mu_3}{\sigma^3}$ є нормований момент третього порядку.

Про наявність асиметрії у досліджуваному розподілі можна судити також по неспівпаданню показників центра розподілу – середнього арифметичного і моди: чим більша різниця $\bar{x} - Mo$, тим більша асиметрія розподілу.

За наявності додатної (правосторонньої) скісності (права гілка кривої довша за ліву) між показниками центра розподілу існує співвідношення: $Mo < Me < \bar{x}$. За наявності лівосторонньої скісності $Mo > Me > \bar{x}$. В симетричному ряду розподілу *As=0*, при правосторонній скісності *As>0*, при лівосторонній – *As<0*.

Для характеристики ступеня гостровершинності використовують четвертий нормований центральний момент. У нормальному розподілі справедливе співвідношення: $\mu_4 = 3\sigma^2$ або $\mu_4 = 3\mu_2$. Таким чином, для нормального розподілу четвертий момент становить $\frac{\mu_4}{\sigma^4} = 3$. Це значення можна використовувати як точку відліку при оцінці міри гостровершинності. Показник

гостровершинності (ексцес) розраховується як різниця між четвертим нормованим моментом для досліджуваного і для нормального розподілів:

$$Ex = \frac{\mu_4}{\sigma^4} - 3.$$

При нормальному розподілі $Ex=0$, при гостровершинному (додатному) ексцесі $Ex>0$, при плосковершинному (від'ємному) – $Ex<0$.

## 5.2 Розрахунок показників

При розрахунку показників можливі два підходи. Перший з них ґрунтується на використанні статистичних даних, котрі після їх систематизації за певною ознакою дозволяють розрахувати необхідні показники. За відсутності статистичних даних застосовується інший підхід, оснований на використанні математичної моделі, в котру входять задані евристично функціональні залежності. На основі цих залежностей проводяться розрахунки.

Перший підхід розглянемо на наступному прикладі.

Статистика інформаційних нападів зображена в таблиці, де приведені значення: $i$ – частка вилученої інформації в $j$-му номері варіаційного ряду, $n_j$ – кількість нападів, що відповідають значенню $i_j$, $n = \sum_j n_j$ – загальна кількість нападів.

Таблиця 5.1

Приклад варіаційного ряду

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $i_j$ | 0 | 0,02 | 0,04 | 0,06 | 0,08 | 0,10 | 0,12 |
| $n_j$ | 1 | 2 | 3 | 4 | 5 | 3 | 1 |

Гістограма варіаційного ряду зображена на рис. 5.1.

За даним (дискретним) статистичним розподілом вибірки необхідно обчислити:

- вибіркову середню величину $\bar{i}$ розподілу;
- дисперсію $D$;
- середнє квадратичне відхилення $\sigma$;

- моду *Mo*;
- медіану *Me*;
- розмах варіації *R*;
- коефіцієнт варіації *V*;
- коефіцієнт асиметрії *As*;
- ексцес *Ex*.

Проведемо розрахунки. Середня арифметична зважена:

$$\bar{i} = \frac{\sum_{j=1}^{7} i_j n_j}{n} = \frac{0 \cdot 1 + 0{,}02 \cdot 2 + 0{,}04 \cdot 3 + 0{,}06 \cdot 4 + 0{,}08 \cdot 5}{1+2+3+4+5+3+1} +$$

$$+ \frac{0{,}10 \cdot 3 + 0{,}12 \cdot 1}{1+2+3+4+5+3+1} = \frac{1{,}22}{19} = 0{,}064.$$

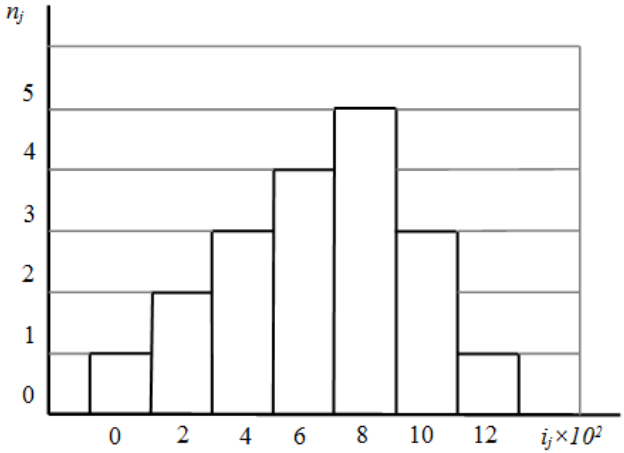



Рис. 5.1. Гістограма розподілу

Дисперсія:

$$D = \frac{\sum_{j=1}^{i_{max}} (i_j - \bar{i})^2 n_j}{n} = \frac{(0-0{,}064)^2 \cdot 1 + (0{,}02-0{,}064)^2 \cdot 2}{19} +$$

$$+ \frac{(0{,}04-0{,}064)^2 \cdot 3 + (0{,}06-0{,}064)^2 \cdot 4 + (0{,}08-0{,}064)^2 \cdot 5}{19} +$$

$$+ \frac{(0{,}10-0{,}064)^2 \cdot 3 + (0{,}12-0{,}064)^2 \cdot 1}{19} = \frac{24{,}4 \cdot 10^{-3}}{19} = 1{,}28 \cdot 10^{-3}$$

Середнє квадратичне відхилення: $\sigma = \sqrt{12{,}8 \cdot 10^{-4}} = 3{,}6 \cdot 10^{-2}$.

Мода: $\sigma = \sqrt{12{,}8 \cdot 10^{-4}} = 3{,}6 \cdot 10^{-2}$.

Медіана: $Mo = 0{,}08$.

Розмах варіації: $R = i_{max} - i_{min} = 12 - 0 = 12$.

Коефіцієнт варіації: $V = \frac{\sigma}{\bar{i}} = \frac{0{,}036}{0{,}064} = 0{,}56$.

Центральний момент третього порядку:

$$\mu_3 = \frac{\sum_{j=1}^{7} (i_j - \bar{i})^3 n_j}{n} = \frac{(0-0{,}064)^3 \cdot 1 + (0{,}02-0{,}064)^3 \cdot 2}{19} +$$

$$+ \frac{(0{,}04-0{,}064)^3 \cdot 3 + (0{,}06-0{,}064)^3 \cdot 4 + (0{,}08-0{,}064)^3 \cdot 5}{19} +$$

$$+ \frac{(0{,}10-0{,}064)^3 \cdot 3 + (0{,}12-0{,}064)^3 \cdot 1}{19} = -9 \cdot 10^{-6}$$

Коефіцієнт асиметрії: $As = \frac{\mu_3}{\sigma^3} = \frac{-9 \cdot 10^{-6}}{46{,}6 \cdot 10^{-6}} = -0{,}19$

Коефіцієнт асиметрії невеликий за абсолютним значенням і від'ємний за знаком. Це значить, що ми маємо невелику лівосторонню скісність.

Центральний момент четвертого порядку:

$$\mu_4 = \frac{\sum_{j=1}^{7}(i_j - \bar{i})^4 n_j}{n} = \frac{(0-0{,}064)^4 \cdot 1 + (0{,}02-0{,}064)^4 \cdot 2}{19} +$$
$$+ \frac{(0{,}04-0{,}064)^4 \cdot 3 + (0{,}06-0{,}064)^4 \cdot 4 + (0{,}08-0{,}064)^4 \cdot 5}{19} +$$
$$+ \frac{(0{,}10-0{,}064)^4 \cdot 3 + (0{,}12-0{,}064)^4 \cdot 1}{19} = 2{,}1 \cdot 10^{-6}$$

Ексцес: $Ex = \frac{\mu_4}{\sigma^4} - 3 = -2{,}76$

Оскільки *Ex* має досить велике значення і від'ємний знак, то закон розподілу є чітко вираженим туповершинним (а не гостровершинним – при *Ex>0*). Розраховані ознаки можна спостерігати на гістограмі (рис. 5.1).

Розглянемо тепер другий підхід. В найпростішому випадку система містить один об'єкт, і цільова функція має вигляд

$$i(x, y) = q(x, y) \cdot f(x, y) ,$$

де *q(x,y)* – щільність розподілу імовірностей *q(x)* при заданому значенні *y*. За цих умов ліва частина рівності також виражає кількість розподілу *i(x)*. Для того, щоб мати можливість розрахувати значення *i(x)* в окремих точках, необхідно перейти до розгляду елементарних подій, тобто значення *q(x)* розглядати як усереднене по малому інтервалу *Δx*:

$$q(x_1 < x < x_2) = \int_{x_1}^{x_2} q(x)dx ,$$

де $x_1, x_2$ – межі інтервалу *Δx*. При більш спрощеному підході розподіл *q(x)* можна замінити на його найбільш сподіване значення. Задамо *q(x)* у вигляді розподілу Максвела $q(x) = Nx^2 e^{-h^2 x^2}$, де $N = \frac{4h^3}{\sqrt{\pi}}$ – нормовочний множник, і знайдемо математичне сподівання випадкової величини *x*, що відповідає розподілу *q(x)*:

$$M(x)=\int_0^{\infty} x\cdot q(x)dx=\frac{4h^3}{\sqrt{\pi}}\int_0^{\infty} x^3e^{-h^2x^2}dx=\frac{4h^3}{\sqrt{\pi}}\cdot\frac{1}{2h^4}=\frac{2}{h\sqrt{\pi}}.$$

Врахуємо, що $h=\frac{1}{x_m}$, де $x_m$ – значення $x$, що відповідає максимуму залежності $q(x)$ і задається на основі експертної оцінки при побудові математичної моделі.

За цих умов маємо цільову функцію у вигляді:

$$i(x)=\frac{2x_m}{\sqrt{\pi}}\cdot f(x).$$

Якщо $f(x)$ має вигляд степеневої функції, то одержуємо:

$$i(x)=\frac{2x_m}{\sqrt{\pi}}\cdot\frac{x^n}{x^n+c}.$$

Якщо ж $q(x)$ задається у вигляді розподілу Релея $q(x)=Nxe^{-h^2x^2}$, де $N=2h^2$, то $M(x)=2h^2\int_0^{\infty} x^2e^{-h^2x^2}dx=\frac{\sqrt{\pi}}{2h}$, де $h=\frac{1}{x_m\sqrt{2}}$.

Цільова функція має вигляд: $i(x)=\frac{x_m\sqrt{\pi}}{\sqrt{2}}\cdot\frac{x^n}{x^n+c}$.

При заміні розподілу $q(x)$ фіксованим значенням $M(x)$ слід враховувати розкиданість можливих значень. При використанні розподілу Максвела дисперсія випадкових величин $x$ становить:

$$D(x)=\int_{-\infty}^{\infty} x^2q(x)dx-(M(x))^2=2h^2\int_0^{\infty} x^3e^{-h^2x^2}dx-\left(\frac{\sqrt{\pi}}{2h}\right)^2=$$

$$=\frac{1}{h^2}-\frac{\pi}{4h^2}=0{,}43x_m^2.$$

Середньоквадратичне відхилення:

$$\sigma(x)=\sqrt{D(x)}=0{,}65x_m.$$

Це значення дозволяє оцінити рівень ризику втрати інформації.

### Контрольні питання

1. Показники статистичних розподілів, їх визначення і взаємозв'язок.
2. Моменти розподілу, їх використання при розрахунку показників.
3. Форма варіаційного ряду і побудова гістограми розподілу.
4. Розрахунок дисперсії випадкових величин $x$ при використанні розподілів Максвела і Релея.

## VI. ТЕОРІЯ ІГОР В ЕКОНОМІЧНИХ ЗАДАЧАХ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

### 6. 1 Матричні ігри. Чисті стратегії.

Захист інформації і розробка засобів протидії спробам її вилучення протилежною стороною є типовий приклад протистояння, яке спостерігається в різних областях людської діяльності. Дослідження цих ситуацій привело до появи нового розділу математики – теорії ігор, яка займається математичним моделюванням ситуацій і розробкою методів пошуку оптимальних рішень. Відповідно до цієї теорії протистояння в інформаційній сфері класифікуються як несиметрична гра з нульовою сумою, оскільки виграти може тільки перший гравець (напад), причому його виграш – кількість вилученої інформації – дорівнює програшу другого гравця (захисту). Метою аналізу є пошук оптимальних рішень для кожної з сторін.

Розглянемо систему, яка складається з $l$ об'єктів. Позначимо розподіл ресурсів нападу через

$$\bar{x} = \{x_1, x_2, ..., x_l\},$$

де $\sum_{k=1}^{l} x_k = X$, $x_k \geq 0$, $k$ – номер об'єкта, $x_k$ – кількість ресурсів нападу, направлених на $k$-ий об'єкт. Розподіл ресурсів захисту:

$$\bar{y} = \{y_1, y_2, ..., y_l\},$$

де $\sum_{k=1}^{l} y_k = Y, y_k$ – кількість ресурсів захисту, направлених на *k*-ий об'єкт. Розподіл $\bar{x}$ і $\bar{y}$ визначають стратегії нападу і захисту.

Для ілюстрації метода аналізу використаємо більш просту модель Гроса, в якій цільова функція, яка виражає частку вилученої інформації, є кусочно-лінійною функцією $x_k$і $y_k$:

$$i(x,y) = \sum_{k=1}^{l} i_k(x,y) = \sum_{k=1}^{l} g_k(x_k - y_k)$$

Побудуємо платіжну матрицю:

$$i = \begin{pmatrix} i_{11} & i_{12} & \dots & i_{1n} \\ i_{21} & i_{22} & \dots & i_{2n} \\ \dots & \dots & \dots & \dots \\ i_{m1} & i_{m2} & \dots & i_{mn} \end{pmatrix}$$

В цій матриці рядки відповідають стратегіям нападу, стовпчики – стратегіям захисту. Елемент $i_{ij}$визначає частку вилученої інформації з системи при *i*-му варіанті розподілу ресурсів нападу по об'єктах ( $i = \overline{1,m}$ )і *j*-му варіанті розподілу ресурсів захисту ( $j = \overline{1,n}$ ). Якщо гравець використовує обрані стратегії з стовідсотковою імовірністю, то такі стратегії називають чистими. Ми будемо розглядати саме такі стратегії.

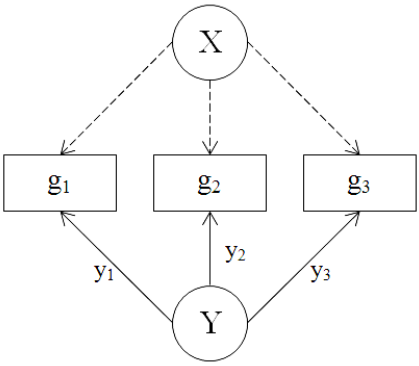



Рис.6.1 Схема протистояння у випадку, коли напад здійснюється на один з об'єктів

Поставимо себе на місце служби захисту інформації. Тоді нашою метою буде мінімізація функції *i(x,y)*, а завданням – вибір розподілу *{$y_k$}*, який забезпечить досягнення $i_{min}$при всіх можливих варіантах розподілу *{$x_k$}*. Фактично ми стоїмо перед необхідністю рішення зворотної задачі: по умові, накладеній на функцію *i(x,y)* потрібно

знайти значення незалежних змінних. Через складність такі задачі часто зводять до прямих і розв'язок знаходять методом перебору.

В наших умовах задача зводиться до розрахунку значень $i(x,y)$ при різних варіантах розподілу $\{x_k\}$ і $\{y_k\}$ з подальшим вибором оптимального варіанта.

Спрощуючи задачу розглянемо випадок, коли ресурси нападу зосереджені на якомусь одному з об'єктів, а ресурси захисту розподілені між об'єктами. Схема протистояння зображена на рис. 6.1. В цьому випадку метою захисту є мінімізація інформації, яку можна вилучити з об'єкта. Цільова функція приймає вигляд:

$$i(x,y) = i_k(x,y) = g_k(X - y_k).$$

Таблиця 6.1

Результат розрахунків функції $i_{jk}(x,y)$ при різних розподілах $\{y_{jk}\}$

| k / j | | 1 | 2 | 3 | $\beta_j$ |
|---|---|---|---|---|---|
| 1 | $y_{1k}$<br>$i_{1k}$ | 0.2<br>0.16 | 0.3<br>0.21 | 0.5<br>0.25 | 0.25 |
| 2 | $y_{2k}$<br>$i_{2k}$ | 0.1<br>0.18 | 0.3<br>0.21 | 0.6<br>0.2 | 0.21 |
| 3 | $y_{3k}$<br>$i_{3k}$ | 0.03<br>0.194 | 0.35<br>0.195 | 0.62<br>0.19 | 0.195 |
| 4 | $y_{4k}$<br>$i_{4k}$ | 0<br>0.2 | 0.35<br>0.195 | 0.65<br>0.175 | 0.2 |
| 5 | $y_{5k}$<br>$i_{5k}$ | 0.1<br>0.18 | 0.2<br>0.24 | 0.7<br>0.15 | 0.24 |
| 6 | $y_{6k}$<br>$i_{6k}$ | 0<br>0.2 | 0.1<br>0.27 | 0.9<br>0.05 | 0.27 |
| 7 | $y_{7k}$<br>$i_{7k}$ | 0<br>0.2 | 0<br>0.3 | 1<br>0 | 0.3 |
| $\alpha_k$ | | 0.16 | 0.195 | 0.15 | |

Таблиця 6.2

Платіжна матриця

| k / j | 1 | 2 | 3 | $\beta_j$ |
|---|---|---|---|---|
| 1 | 0.16 | 0.21 | 0.25 | 0.25 |
| 2 | 0.18 | 0.21 | 0.2 | 0.21 |
| 3 | 0.194 | 0.195 | 0.19 | 0.195 |
| 4 | 0.2 | 0.195 | 0.175 | 0.2 |
| 5 | 0.18 | 0.24 | 0.15 | 0.24 |
| 6 | 0.2 | 0.27 | 0.05 | 0.27 |
| 7 | 0.2 | 0.3 | 0 | 0.3 |
| $\alpha_k$ | 0.16 | 0.195 | 0 | |

Розглянемо приклад: покладемо $X=Y=1$, $l=3$, $g=g_1+g_2+g_3=1$, $g_1=0.2$, $g_2=0.3$, $g_3=0.5$.Результати розрахунків функції

$$i_{jk}(x,y)=g_k(1-y_{jk})$$

для різних варіантів розподілу$y_{jk}$( $j=\overline{1,7}$ ) зображені в табл.6.1. На основі цих даних побудована платіжна матриця (табл. 6.2), в якій приведені також значення $\beta_j=\max\limits_k i_{jk}$ і $\alpha_k=\max\limits_j i_{jk}$. Число $\beta_j$називають показником неефективності стратегії захисту (оскільки воно представляє максимальне значення вилученої інформації), а $\alpha_k$ – показником ефективності стратегії нападу.

Вибір варіантів розподілу $\{y_{jk}\}$ приводимо з таких міркувань. Перший варіант повторює розподіл $\{g_k\}$ – 0.2:0.3:0.5. В подальшому ми переміщуємо ресурси в напрямку об’єкта, на якому значення $i_{jk}$ перевищує інші. Рівноважний варіант досягається при такому розподілі $\{y_{jk}\}$, коли всі $i_{jk}$однакові (в такому випадку це третій варіант $j=3$). В цьому варіанті досягається $\min\limits_j \alpha_j=\min\limits_j \max\limits_k i_{jk}=0.195$ при розподілі $y_1:y_2:y_3=0.03:0.35:0.62$. Цей варіант гарантує, що частка вилученої інформації не перевищить значення 0.195 при будь-яких варіантах нападу на об’єкти.

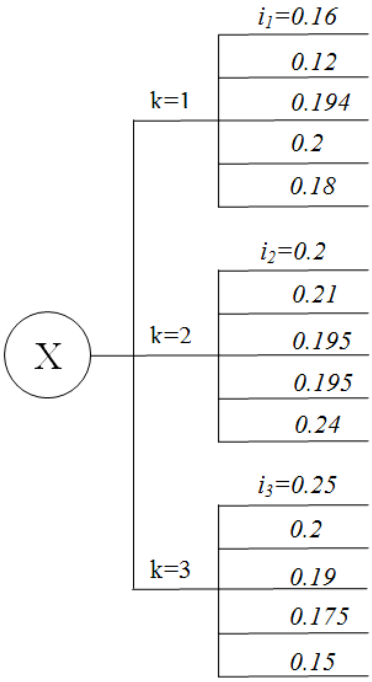



Рис.6.2 Дерево рішень для сторонни нападу

Для нападу оптимальним є розподіл $x_1:x_2:x_3=0:1:0$–всі ресурси вкладаються в другий об'єкт. В цьому варіанті частка вилученої інформації буде не менша 0.195 при всіх варіантах розподілу ресурсів захисту. Таким чином, другий стовпчик і третій рядок представляють оптимальні чисті стратегії для першого (напад) і другого (захист) гравців. Оптимальні значення функції для обох сторін $\alpha = \max_k \min_j i_{jk}$ і $\beta = \min_j \max_k i_{jk}$ співпадають і становлять ціну гри $\alpha=\beta=v=0.195$. В геометричній інтерпретації одержаний результат зображується сідловою точкою на гіперповерхні $i(x,y)$. Ситуація, яка відповідає сідловій точці, є рівноважною, оскільки ні одна зі сторін не зацікавлена в тому, щоб її порушити.

У випадку, коли сторони мають змогу приймати альтернативні рішення, розв'язок наочно можна представити у вигляді дерева рішень. Для нападу це дерево зображене на рис. 6.2.

## 6.2 Критерії оптимальності. Змішані стратегії.

Зауважимо, що застосований критерій (він називається критерієм мінімакса або критерієм Севіджа) не є єдино можливим. Його привабливість в задачах інформаційної безпеки обумовлена тим, що одержаний розподіл $\{y_k\}$ гарантує мінімальний ризик перевищення відповідного значення $i(x,y)$. Тому цей критерій називають ще критерієм обережності. Він характерний для людей не схильних до ризику.

Узагальненням критерію Севіджа можна вважати критерій Гурвіца, за яким вибір рішення направлений на досягнення деякого середнього значення ефективності, котре враховує як максимальне, так і мінімальне значення ефективності для кожного варіанта. Кожне з цих значень враховується при усередненні з певним ваговим коефіцієнтом $\lambda$, який характеризує схильність менеджера до ризику: для схильних до ризику $\lambda \to 1$, для не схильних – $\lambda \to 0$, для нейтральних – $\lambda=0,5$. За цим критерієм середньозважена кількість вилученої інформації знаходиться за виразом:

$$\overline{i_k} = \lambda i_{k\,max} + (1-\lambda) i_{k\,min}$$

Поклавши для прикладу $\lambda=0,5$ з табл.2 знайдемо для кожного з об'єктів:

$$\overline{i_1} = 0.5(0.2 + 0.16) = 0.18;$$

$$\overline{i_2} = 0.5(0.3 + 0.195) = 0.248;$$

$$\overline{i_3} = 0.5(0.25 + 0) = 0.125.$$

За критерієм Гурвіца при $\lambda=0,5$ напад слід здійснювати також на другий об'єкт: $\overline{i_{k\,max}} = \overline{i_2} = 0.125$, але очікувана частка вилученої інформації зросла порівняно з даними табл.2, оскільки враховується можливість досягнення максимального значення $i$.

Проведені розрахунки можна продовжити при інших значеннях параметрів і більш досконалій постановці задачі. Зокрема, можна розглянути варіанти з різними розподілами $\{g_k\}$, співвідношеннями $Z = \frac{X}{Y}$, з розподілом ресурсів нападу $\{x_k\}$ між всіма об'єктами, іншими критеріями оптимальності і багатоцільовій функції.

В нашому випадку припускалось, що імовірності нападів на різні об'єкти однакові. Це відповідає ситуації, коли напад не має відомостей про кількість інформації на об'єктах і стан їх захищеності. У випадку певної обізнаності слід враховувати імовірності нападів, і задача стає стохастичною. При цьому ми переходимо від стану невизначеності до ситуації, коли рішення приймається в умовах ризику. В цій ситуації можна використати критерій Бейєса, за яким показником ефективності вкладання

коштів є середнє значення або математичне сподівання виграшу з врахуванням ймовірностей всіх можливих станів. Таким чином, обране за цим критерієм рішення є оптимальним не в кожному окремому випадку, як для критерію Севіджа, а в середньому.

В табл.6.3 приведені усереднені по об'єктах значення і для двох випадків: середнє арифметичне значення $\overline{i_j}$ (імовірності нападів на об'єкти однакові $p_1=p_2=p_3=0.33$) і математичне сподівання $\overline{i_j}(p)$, одержане при $p_1=0.2$, $p_2=0.3$, $p_3=0.5$.

Таблиця 6.3

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $\overline{i_j}$ | 0.207 | 0.197 | 0.193 | 0.19 | 0.19 | 0.173 | 0.167 |
| $\overline{i_j}(p)$ | 0.22 | 0.199 | 0.193 | 0.187 | 0.183 | 0.146 | 0.13 |

Обидві усереднені величини приймають свої мінімальні значення у випадку $j=7$, коли всі ресурси захисту зосереджені на третьому, найважливішому об'єкті, а не розподілені між об'єктами, як це було при використанні критерію Севіджа. Встановлення умов, за яких слід переходити від одного принципу розподілу до другого, є одним з важливих завдань аналізу.

Можливі випадки, значення $\alpha = \max_k \min_j i_{jk}$ і $\beta = \min_j \max_k i_{jk}$ не співпадають, при чому $\beta>\alpha$. Величину $\alpha$ називають нижньою, а $\beta$ – верхньою ціною гри. Сідлова точка відсутня, тому застосування чистих стратегій не дає оптимального рішення. В цьому випадку застосовують змішані стратегії для кожної сторони. Змішана стратегія полягає у випадковому виборі окремих чистих стратегій з певними ймовірностями $p_1,p_2,\ldots p_n$, $\sum_{j=1}^{n} p_n = 1$. Змішану стратегію, яка містить $n$ чистих стратегій, можна представити $n$-мірним вектором $P=(p_1,p_2,\ldots p_n)$. Знаходження значень $p_1,p_2,\ldots p_n$ і є нашим завданням при застосуванні змішаної стратегії. В нашій задачі змішана стратегія захисту полягає в застосуванні розподілів ресурсів захисту, які визначаються номером $j$ (табл.6.1) з певними ймовірностями, змішана стратегія нападу – у випадковому виборі об'єктів нападу ($k=\overline{1,3}$) зі своїм набором ймовірностей.

Зрозуміло, що змішані стратегії застосовуються при багаторазовому повторенні гри.

### Контрольні питання

1. Принципи застосування теорії ігор до рішення задач інформаційної безпеки.
2. Представлення розв'язку оптимізаційних задач у вигляді дерева рішень.
3. Критерії оптимальності, їх порівняльна характеристика і сфера застосування.
4. Чисті і змішані стратегії, їх сутність.

## VII. МЕТОДИ РІШЕННЯ ОПТИМІЗАЦІЙНИХ ЗАДАЧ

### 7.1 Класичні методи оптимізації

При дослідженні операцій частіше за все необхідно знайти максимум або мінімум певної величини. Ці два поняття об'єднуються терміном *екстремум*. Визначимо умови досягнення екстремуму цільової функції.

**Теорема Вейєрштраса:** Якщо область $D$ замкнута і обмежена, то диференційована функція $z=f(X)$, $X \in D$ досягає в цій області глобального екстремуму (найбільшого і найменшого значення)в стаціонарній або в граничній точці області.

*Стаціонарною точкою* $X^0$ називається точка, в якій всі частинні похідні функції$z=f(X)$ дорівнюють нулю.

Розглянемо спочатку одновимірний простір, в якому$X=x$і $z=f(X)$є функція однієї змінної.

Маємо різні форми залежності $z=f(X)$.

1) Лінійна функція $z=kx+b$

Стаціонарних точок лінійна функція не має і досягає свого найменшого і найбільшого значення в

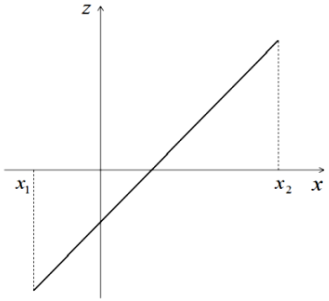



Рис.7.1 Лінійна функція

граничних точках $x_1$ і $x_2$ відповідно (рис. 7.1).

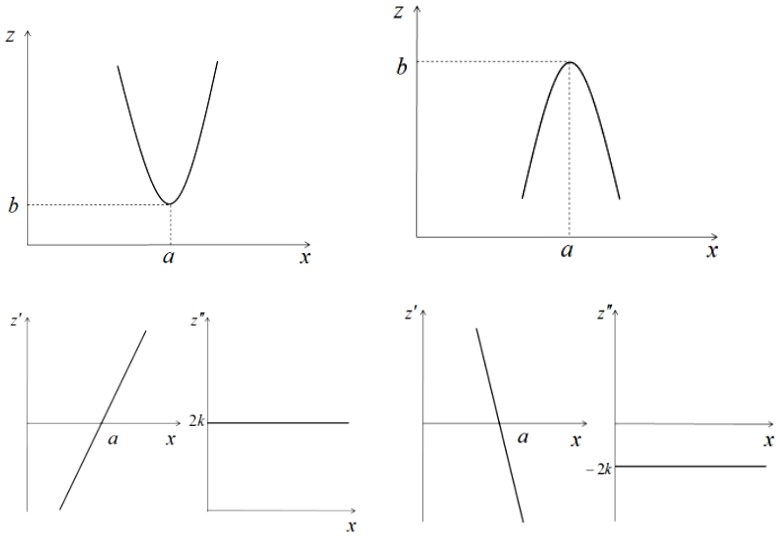
2) Квадратична функція

а) $z=k(x-a)^2+b$

Екстремум функції $z_{min}=b$ визначається з умови$z'=2k(x-a)=0$,і знаходиться в точці $x^0=a$ (рис.7.2,а). Вид екстремуму визначається з умови $z''=2k>0$ (мінімум).

б) $z=-k(x-a)^2+b;$ $z'=-2k(x-a)$

В цьому випадку $z''=-2k<0$, що визначає вид екстремуму




а) б)

Рис.7.2 Квадратичні функції і їх похідні

(максимум) в точці $x^0=a$.

Розглянемо тепер функцію двох змінних: $z=f(x_1,x_2)$. Тут можливі такі варіанти. Якщо в стаціонарній точці $(x_1^0,x_2^0)$ по обом координатам маємо *min*, то $z=f(x_1,x_2)$ в цій точці також має *min*, а якщо по обом координатам спостерігається max, то – *max*. Якщо хоча б по одній з координат ця точка є точкою перетину, то екстремум відсутній. Якщо ж по одній з координат маємо *min*, а по іншій *max*, то одержуємо сідлову точку (рис.7.3).

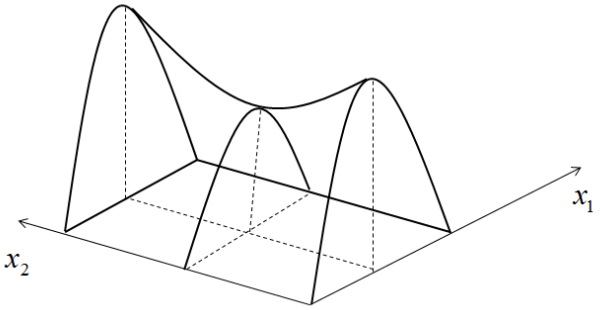



Рис.7.3 Утворення сідлової точки

Отже, якщо функція по обом координатам має квадратичну форму (в загальному випадку – степеневу форму парного ступеня), то маємо екстремум або сідлову точку, а якщо кубічну форму (степеневу форму непарного ступеня), то екстремуму нема.

Сформулюємо необхідні і достатні умови екстремуму функції багатьох змінних в загальному вигляді.

Необхідна умова: в точці екстремуму частинні похідні по незалежним змінним дорівнюють нулю:

$$z'_{x_i}(x^0) = 0$$

Для двох змінних:

$$\begin{cases} \left.\dfrac{\partial z}{\partial x_1}\right|_{x^0} = 0 \\ \left.\dfrac{\partial z}{\partial x_2}\right|_{x^0} = 0 \end{cases}$$

Для одержання достатніх умов необхідно визначити в стаціонарній точці знак диференціала другого порядку. Позначимо:

$$a_{11} = \left.\frac{\partial^2 z}{\partial x_1^2}\right|_{x^0}; \qquad a_{12} = \left.\frac{\partial^2 z}{\partial x_1 \partial x_2}\right|_{x^0};$$

$$a_{22} = \left.\frac{\partial^2 z}{\partial x_2^2}\right|_{x^0}; \qquad a_{21} = \left.\frac{\partial^2 z}{\partial x_2 \partial x_1}\right|_{x^0}.$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Зазначимо, що якщо функції $a_{12}$ і $a_{21}$ неперервні, то $a_{21}=a_{12}$.

Достатні умови екстремуму функції двох змінних мають вигляд:

А) якщо $\Delta>0$, то при $a_{11}<0$ $(a_{22}<0)$ маємо *max;*

при $a_{11}>0$ $(a_{22}>0)$ - *min.*

Б) якщо $\Delta<0$, то екстремуму немає.

В) якщо $\Delta=0$, то маємо сідлову точку.

В оптимізаційних задачах часто разом із функцією, для котрої необхідно знайти екстремум, задають обмеження на параметри і змінні величини, які входять в цю функцію. Екстремум, досягнутий за умови виконання таких обмежень, називається умовним. В економічних задачах інформаційної безпеки обмеження зазвичай стосуються ресурсів нападу і захисту, які повинні задовольняти рівностям:

$$\sum_{k=1}^{n} x_k = X; \quad \sum_{k=1}^{n} y_k = Y.$$

Приведені вирази носять назву рівнянь зв'язку.

В деяких випадках рівності замінюють на нерівності. При пошуку рішення обмеження можуть мати суттєве значення, оскільки оптимальний розподіл ресурсів може досягатись за такої їх кількості, яка виходить за межі допустимих значень.

*Приклад.1* Маємо систему з двох об'єктів. Допустимі значення ресурсів двох сторін становлять $x_{гр}=x_1+x_2=6$, $y_{гр}=y_1+y_2=4$. Цільові функції нападу і захисту задаються у вигляді квадратичних залежностей від двох змінних.

1. Цільова функція нападу:

$$i(x_1,x_2)=-x_1^2+6x_1-2x_2^2+8x_2-9= \tag{1}$$

$$=-(x_1^2-2x_1\cdot 3+3^2-3^2)-2(x_2^2-2x_2\cdot 2+2^2-2^2)=-(x_1-3)^2+9-2(x_2-2)^2+8-9=-(x_1-3)^2-2(x_2-2)^2+8\rightarrow max$$

Визначимо вид екстремуму в стаціонарній точці А(3;2):

$$\frac{\partial i}{\partial x_1} = -2(x_1-3); \qquad \frac{\partial i}{\partial x_2} = -4(x_2-2);$$

$$a_{11} = \frac{\partial^2 i}{\partial x_1^2} = -2\ ; \qquad a_{12} = \frac{\partial^2 i}{\partial x_1 \partial x_2} = 0\ ;$$

$$a_{21} = \frac{\partial^2 i}{\partial x_2 \partial x_1} = 0\ ; \qquad a_{22} = \frac{\partial^2 i}{\partial x_2^2} = -4\ .$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ 0 & -4 \end{vmatrix} = 8.$$

Маємо $\Delta>0$, $a_{11}<0$, отже функція $i(x_1,x_2)$ в стаціонарній точці A(3;2) має максимум (рис.7.4 )

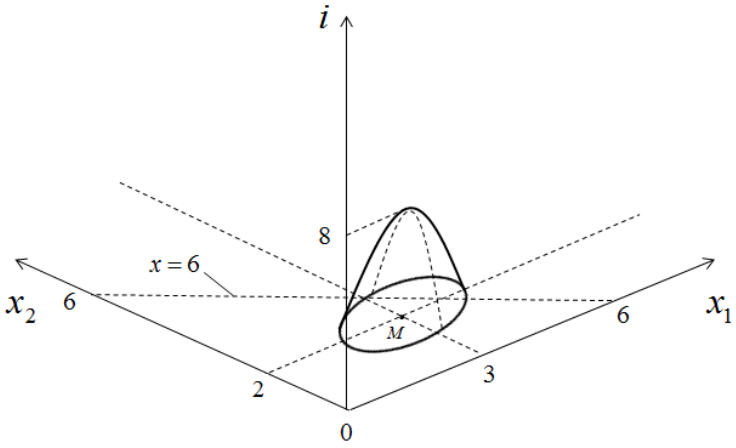



Рис.7.4 Максимум функцій двох змінних

Цільова функція захисту:

$$i(y_1,y_2)=2y_1^2-4y_1+y_2^2-6y_2+13=$$
$$=2(y_1^2-2y_1\cdot 1+1-1)+y_2^2-2y_2\cdot 3+3^2-3^2+13= \qquad (2)$$
$$=2(y_1-1)^2+(y_2-3)^2+2$$

Знайдемо вид екстремуму в стаціонарній точці B(1;3):

$$\frac{\partial i}{\partial y_1} = 4(y_1 - 1)\ ; \qquad \frac{\partial i}{\partial y_2} = 2(y_2 - 3)\ ;$$

$$a_{11} = \frac{\partial^2 i}{\partial y_1^2} = 4\ ; \qquad a_{22} = \frac{\partial^2 i}{\partial y_2^2} = 2\ ; \qquad a_{12} = a_{21} = 0;$$

$$\Delta = \begin{vmatrix} 4 & 0 \\ 0 & 2 \end{vmatrix} = 8.$$

Оскільки $\Delta>0$ і $a_{11}>0$, то маємо мінімум (рис.7.5)

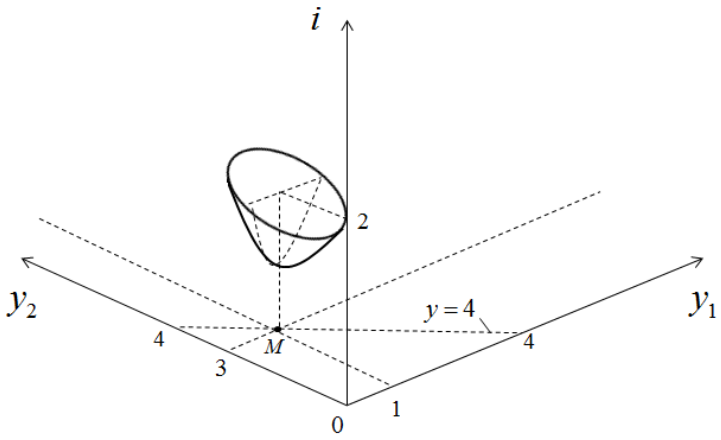



Рис.7.5 Мінімум функції двох змінних

Функція $i(x_1,x_2)$ описує дії нападу та визначає оптимальний розподіл його ресурсів $\{x_1^0,x_2^0\}=\{3;2\}$ і відповідну кількість вилученої інформації $i_{max}=8$. Функція $i(y_1,y_2)$ визначає оптимальний розподіл ресурсів захисту $\{y_1^0,y_2^0\}=\{1;3\}$, для якого маємо $i_{min}=2$.

Числові значення, які входять у вирази в дужках, а також вільні члени в (*1*) і (*2*) віднесені до загальної кількості інформації і виражені у відсотках. Так з (*1*) випливає, що найбільша кількість вилученої інформації становить 8% і досягається при розподілі загальної кількості ресурсів нападу $X=x_1+x_2=5$ у такій пропорції: на перший об'єкт виділяється 3% від вартості всієї інформації, а на другий – 2% (для спрощення вважаємо кількість інформації на об'єктах однаковою).

Баланс витрат і доходів можна прокоментувати наступним чином. Витрати нападу $x=5$ не перевищують допустимої величини $x_{гр}=6$, а дохід $i_{max}=8$ перевищує витрати. Витрати захисту $y=4$ знаходяться в межах допустимих значень $y<y_{гр}$, а дохід від

інвестицій становить $j=8-2=6>y$. Таким чином, обидві операції рентабельні.

Обмеження ресурсів при геометричній інтерпретації результатів проявляється в тому, що проекція точки, яка відповідає оптимуму цільової функції на площину $x_1 0 x_2$ чи $y_1 0 y_2$ повинна знаходитись всередині або, принаймні, на гіпотенузі рівнобедреного прямокутного трикутника з катетами X і, відповідно, Y.

Слід зауважити, що кількість вилученої інформації залежить не тільки від розподілу власних ресурсів, а й ресурсів протилежної сторони. В $i(x_1,x_2)$ величини $y_1$, $y_2$ не входять в явному вигляді, проте вони визначають числові параметри в (1) і таким чином опосередковано впливають на шуканий розподіл $\{y_k^0\}$ аналогічно параметрам $n$і $c$в дробно-нелінійних цільових функціях.

В реальних умовах цільові функції мають більш складний характер, проте в багатьох випадках їх геометрична форма близька до приведеної у наших прикладах дзвоноподібної форми.

При динамічному протистоянні, коли кожна з сторін почергово здійснює свої кроки, ми переходимо поступово від рис.7.4 до рис.7.5, потім знов до рис.7.4, в якому з врахуванням скоригованих значень $\{y_k\}$ дещо змінюється кривизна і положення просторової фігури, і так далі – поки вершини обох «дзвонів» не зблизяться на мінімально можливу відстань або не співпадуть. На першому етапі розподіляємо певним чином ресурси захисту (скажімо, пропорційно кількості інформації на об'єктах) і знаходимо оптимальний розподіл ресурсів нападу $\{x_k^0\}$ (рис.7.4). Другий крок: вважаючи розподіл $\{x_k\}$ відомим, знаходимо оптимальний розподіл ресурсів захисту $\{y_k^0\}$ і подаємо його у вигляді залежності $i(y_1,y_2)$ – рис.7.5. При зміні $\{y_k\}$ змінюються значення числових коефіцієнтів в (1), що призводить до зміни форми і положення просторової фігури (рис.6.4) і, зрештою, – значень $\{x_k^0\}$ і $i_{max}$.

## 7.2. Геометричний метод рішення задач лінійного програмування

### *7.2.1. Сутність геометричного методу*

При дослідженні операцій нелінійні процеси, які відбуваються в реальних умовах, часто пробують звести до більш простих лінійних процесів. В результаті ми приходимо до задачі лінійного програмування, сутність якої полягає в наступному.

Маємо систему *m* лінійних рівнянь або нерівностей з *n* змінними:

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \le b_1$$

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \le b_2$$

……

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \le b_m,$$

де $m<n$ (така система називається невизначеною).

Необхідно знайти розв'язок системи:

$$X = (x_1, x_2 ... x_n),$$

де $x_j \ge 0$, $j=1,2..n$ (умова невід'ємності змінних), при якому лінійна форма:

$$F = c_1x_1 + c_2x_2 + .. + c_n.x_n,$$

яку називають цільовою функцією, приймає оптимальне (мінімальне або максимальне) значення. Систему () можна розглядати як обмеження, котрі накладаються на шукану функцію *F*.

Обмежившись розглядом системи рівнянь, загальну задачу лінійного програмування можна представити у вигляді:

$$F = \sum_{j=1}^{n} c_i x_j \to \max \ (\min)$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \ x_j \ge 0, \ j = 1,2...n, \ i = 1,2..m.$$

Задача лінійного програмування відрізняється від стандартного розв'язку системи лінійних алгебраїчних рівнянь тим, що кількість

невідомих перевищує кількість рівнянь і серед можливих розв’язків необхідно знайти такий, який задовольняє умові оптимальності.

У випадку, коли кількість змінних дорівнює двом, розв'язок можна подати в геометричній формі. З лінійної алгебри відомо, що множина допустимих розв’язків задачі лінійного програмування являє собою опуклий многогранник (у випадку двох змінних – многокутник), а оптимальний розв'язок знаходиться принаймні в одній з кутових точок многогранника розв’язків.

Для знаходження розв'язку побудуємо лінії рівня лінійної функції *F*, тобто лінії, на яких функція приймає фіксовані значення*F=a*:

$$c_1 x_1 + c_2 x_2 = a$$

При зміні значення *a* лінії пересувається в системі координат $\{x_1, x_2\}$ паралельно самій собі. Необхідно знайти положення лінії рівня, при якому вона проходить через кутову точку многокутника, яка відповідає максимальному (чи мінімальному) значенню *a*.

Приклад 1. Знайти $\boldsymbol{F = x_1 + 2x_2 \to max}$ при обмеженнях, які включають умови невід'ємності змінних:

$$\begin{cases} x_1 + 3x_2 \le 18 \\ x_1 + x_2 \le 10 \\ x_1 \le 5 \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

Зобразимо многокутник розв’язків і лінії рівня (рис. 7.6). Покладемо *a=4* і побудуємо першу лінію рівня:

$$x_1 + 2x_2 = 4.$$

З теорії поля відомо, що напрямок найшвидшої зміни функції співпадає з перпендикуляром до лінії рівня ( згадаємо, лінії напруженості електростатичного поля перпендикулярні до еквіпотенціальних ліній: $E = -grad\varphi$). Неважко упевнитись, що зростання функції *F* відбувається при віддаленні від початку координат, тобто співпадає з напрямком вектора $\bar{n}$. Переміщуємо лінію рівня, поки вона не співпаде з найбільш віддаленою точкою *C* многокутника. Знайдемо координати цієї точки, розв’язавши систему рівнянь прямих *BC* і *CD*:

$$\begin{cases} x_1 + 3x_2 = 18 \\ x_1 + x_2 = 10, \end{cases}$$

звідки$x_1=6, x_2=4$. Отже, максимальне значення $F=14$ функції $F = x_1 + 2x_2$ при виконанні зазначених обмежень досягається при $x_1=6, x_2=4$.

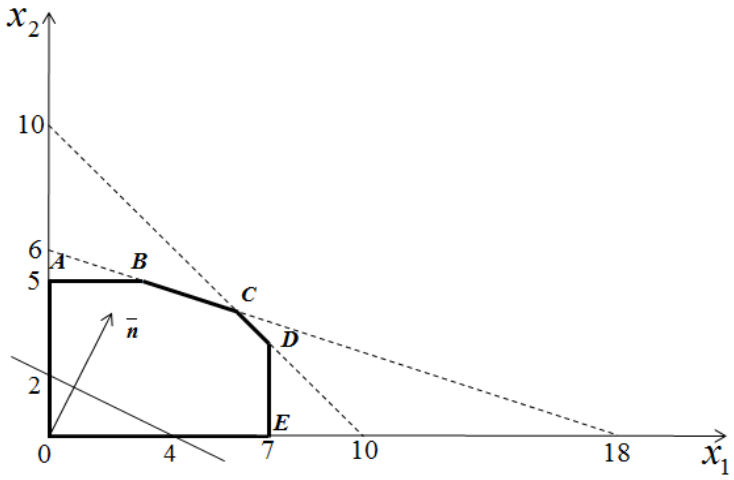



Рис. 7.6. Многокутник допустимих розв'язків для нападу

Спробуємо надати приведеній математичній задачі певного фізичного смислу. Звертаючись до нашої проблематики, вважатимемо, що $x_1$і $x_2$– це ресурси нападу, а $F$ – кількість вилученої інформації, виражена в певних грошових одиницях (г.о.). Отже, розподіливши загальну кількість ресурсів $x_1+ x_2=10$ г.о. між двома об'єктами у пропорції 6:4 при певній кількості ресурсів захисту і заданій вразливості об'єктів (вони визначають коефіцієнти при невідомих у функції $F$) напад вилучить інформації на 14 г.о.

Приклад 2. Знайти $F = 2y_1 + 3y_2 \to min$ при обмеженнях:

$$\begin{cases} 3y_1 + y_2 \geq 9 \\ y_1 + 2y_2 \geq 8 \\ y_1 + 6y_2 \geq 12 \\ y_1 \geq 0,\ y_2 \geq 0. \end{cases}$$

В цій задачі $y_1$ і $y_2$– ресурси захисту, виділені на перший і другий об'єкти, а $F$ - кількість вилученої інформації з об'єктів при заданих кількостях ресурсів нападу і вразливостях об'єктів.

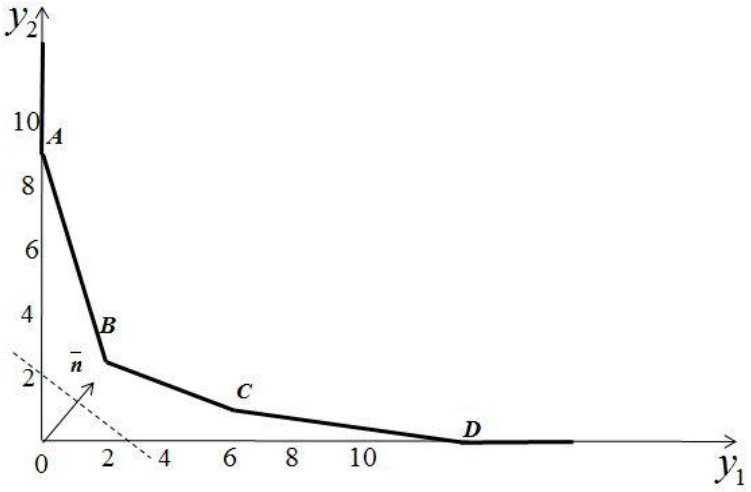



Рис.7.7. Многокутник допустимих розв'язків для захисту

Многокутник розв'язків становить необмежену многокутну область (рис.7.7). Побудуємо первинну лінію рівня $2y_1 + 3y_2 = 6$ і будемо рухати її в бік зростання значень$F$ до дотику лінії з многокутною областю в точці $B$, яка знаходиться на перетині прямих $AB$ і $BC$. Координати цієї точки $y_1=2$, $y_2=3$визначаються з розв'язку системи рівнянь:

$$\begin{cases} 3y_1 + y_2 = 9 \\ y_1 + 2y_2 = 8. \end{cases}$$

При подальшому переміщенні лінії значення $F$ зростатимуть.

Отже, $F_{min}$=13 при загальній кількості ресурсів захисту $y_1 + y_2 = 5$ досягається при розподілі ресурсів між об'єктами у співвідношенні 2:3.

Приклад 3. Знайти $F = 3x_1 + 5x_2 \to max$ при обмеженнях:

$$\begin{cases} 3x_1 - 4x_2 \le 5 \\ 2x_1 \le 7 \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

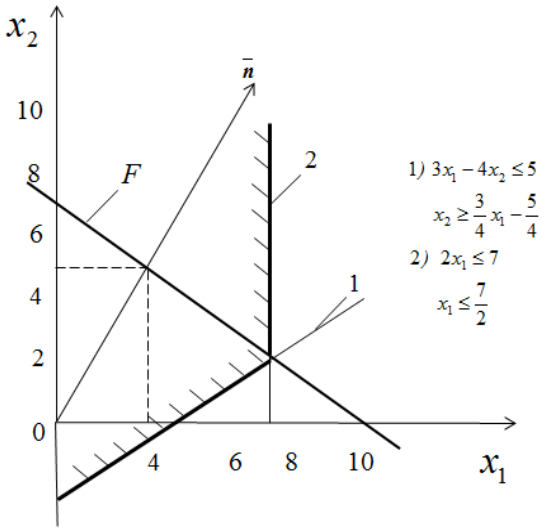



Рис. 7.8. Необмежена область розв'язків.

З рисунка видно, що задача розв'язку не має: $x_2 \to \infty$, бо в умові не поставлена верхня межа для $x_2$.

### *7.2.2 Чутливість рішення до зміни вхідних даних*

В умовах, коли параметри математичної моделі задаються евристично (наприклад, на основі експертної оцінки) важливого значення набуває питання про стійкість рішення, або його

чутливість до зміни вхідних даних. Ці зміни можуть торкатись як коефіцієнтів цільової функції, так і констант, котрі входять в формулювання обмежень. Проілюструємо це на графічному розв'язку прикладу 2. При незначній зміні коефіцієнтів цільової функції буде дещо змінюватись нахил ліній рівня *F=const*, проте вони, як і раніш, будуть торкатись обмежувальної ламаної ABCD в точці B (рис. 7.7). Якщо лінії рівня стануть паралельними відрізку BC, то оптимальні значення будуть відповідати всім точкам цього відрізка, тобто утворять неперервну множину (континуум). При подальшому збільшенні нахилу лінії рівня стануть торкатись обмежувальної ламаної в точці C. Таким чином, можна зробити висновок, що найменш стійкими будуть рішення, при яких нахил ліній рівня наближається до нахилу одного з обмежувальних відрізків. Зокрема, при переході лінії рівня через положення паралельності відрізку BC, спостерігається стрибок оптимального значення з точки B в точку C. При цьому розподіл $\{y_1,y_2\}$, що відповідав точці оптимальності B, буде вже далеким від оптимального. Зроблений висновок накладає певні умови на співвідношення $c_1/c_2$ коефіцієнтів в цільовій функції. Це співвідношення не повинно бути близьким до співвідношень коефіцієнтів в обмежувальних нерівностях (при їх спів падінні лінії будуть паралельними). В нашому прикладі лінія рівня знаходиться за своїм нахилом між лініями AB і BC. Точка B буде залишатись точкою оптимальності, якщо відношення $c_1/c_2$ лежить в межах

$$\frac{1}{2} < \frac{c_1}{c_2} < 3.$$

Якщо один з коефіцієнтів ($c_1$ чи $c_2$) зафіксований, то наведена нерівність дає змогу визначити, в яких межах повинен знаходитись інший коефіцієнт.

## 7.3 Застосування геометричного методу до рішення економічних задач інформаційної безпеки

### *7.3.1 Лінійні задачі*

*Приклад 2.*Визначити оптимальний розподіл ресурсів між трьома об'єктами, якщо прибуток $b_k$ від внесення інвестицій в захист і коефіцієнт ризику $r_k$ для об'єктів визначається даними, приведеними в таблиці. Загальний коефіцієнт ризику повинен задовольняти умові $r \le 1,3$.

| $k$ | $b_k$ | $r_k$ |
|---|---|---|
| 1 | 0,14 | 1,2 |
| 2 | 0,16 | 1,4 |
| 3 | 0,10 | 1,0 |

Задача полягає у визначенні часток $y_1$, $y_2$, $y_3$ інвестицій в кожний об'єкт, які забезпечують максимальний прибуток при заданому граничному рівні $r$ ризику втрати інформації.

Математично задача формулюється наступним чином. Необхідно максимізувати цільову функцію, що визначає загальний прибуток:

$$b(y_1, y_2, y_3) = b_1 y_1 + b_2 y_2 + b_3 y_3 = 0,14 y_1 + 0,16 y_2 + 0,10 y_3 \to max$$

при обмеженнях:

$$\begin{cases} r_1 y_1 + r_2 y_2 + r_3 y_3 = 1,2 y_1 + 1,4 y_2 + 1,0 y_3 \le 1,3 \\ 0 \le y_1 \le 1;\ \ 0 \le y_2 \le 1;\ \ 0 \le y_3 \le 1; \\ y_1 + y_2 + y_3 = 1. \end{cases}$$

Зведемо задачу до двох змінних, виразивши з останнього обмеження $y_3$ через $y_1$ і $y_2$:

$y_3 = 1 - y_1 - y_2$

Тоді цільова функція мети приймає вигляд:

$b(y_1, y_2) = 0,14\, y_1 + 0,16\, y_2 + 0,10(1 - y_1 - y_2) = 0,04\, y_1 + 0,06\, y_2 + 0,10$

при обмеженнях:

$$\begin{cases} 0,2 y_1 + 0,4 y_2 \le 0,3 & (1) \\ 0 \le y_1 \le 1;\ \ 0 \le y_2 \le 1; & (2) \\ 0 \le y_1 + y_2 \le 1. & (3) \end{cases}$$

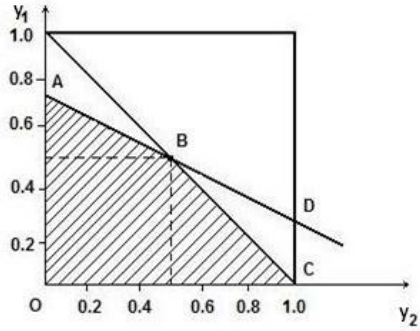



Рис.7.9 Побудова області допустимих розв'язків

В системі координат $y_1Oy_2$ будуємо многокутник OABC, який являється областю допустимих розв'язків системи нерівностей (вона позначена штриховкою на рис.7.9).

Простір допустимих розв'язків формується нерівностями (1) – (3). Кожна з цих нерівностей утворює певну область. Перша з них обмежена прямою *AD*, яка задається рівністю:

$0{,}2\ y_1+0{,}4\ y_2 =0{,}3$

Положення цієї прямої знаходимо, визначивши точки її перетину з осями координат *A (0; 0,75)*, *D (1,5; 0)*. Друга область, яка визначається нерівностями (2), формує квадрат, обмежений прямими $y_1= 0;\ y_1 = 1$ і $y_2 = 0;\ y_2 =1$.Третя область (нерівність (3)) обмежена осями координат і діагоналлю квадрата. Вимога одночасного виконання всіх нерівностей приводить до накладання зазначених областей, що й утворює заштриховану область.

Відомо, що в опуклій області максимум лінійної функції може досягатись лише в одній з її вершин. Знаходимо значення функції $b(y_1,y_2)$ в точках *O, A, B, C*:

$b_O(0; 0) = 0;\ b_A(0; 0{,}75) = 0{,}145;\ b_B(0{,}5; 0{,}5) = 0{,}15;\ b_C(1; 0) = 0{,}14.$

Максимум досягається в точці *B* і становить $b_{max} = 0{,}15$. Перевіримо виконання умови $0{,}2\ y_1+0{,}4\ y_2 \le 0{,}3$в точці *B*:

$0{,}2\ y_1+0{,}4\ y_2 =0{,}2\cdot 0{,}5+0{,}4\cdot 0{,}5 = 0{,}3.$

Отже, умова $r \le 1{,}3$ виконується. Задача розв'язана.

### *7.3.2 Дрібно-лінійні і дрібно-нелінійні задачі*

Важливим напрямком економічного менеджменту інформаційної безпеки є визначення оптимального розміру ресурсів захисту і їх розподілу між об'єктами. У випадку двох об'єктів розв'язок можна подати в геометричній формі. Не зважаючи на обмеження в кількості об'єктів, такий підхід викликає інтерес, оскільки дозволяє наочно продемонструвати методику пошуку оптимуму і його залежність від параметрів математичної моделі.

В умовах невизначеності при пошуку оптимального рішення слід передбачити всі можливі варіанти дій суперника, в тому числі найбільш несприятливий для нас і оптимальний для суперника. Проте при пошуку такого варіанта суперник знаходиться в такому ж стані невизначеності і зазнає таких же труднощів. В результаті ми приходимо до необхідності розв'язку двоїстої задачі шляхом рекурентних процедур, тобто почергового пошуку оптимуму кожної з сторін при прогнозованій стратегії суперника. В термінології теорії ігор це позиційна гра.

Проілюструємо цю методику на прикладі. Розглянемо спочатку дії нападу і використаємо цільову функцію у вигляді [4]:

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

де $i(x,y)$ – частка вилученої інформації;

*x* і *y* – змінні величини, які визначають ресурси нападу і, відповідно, захисту;

$g_k$ – відносна кількість інформації на *k*-му об'єкті;

$p_k(x,y)$ – імовірність нападу на *k*-й об'єкт;

$q_k(x,y)$ – імовірність виділення нападом ресурсів *x* на *k*-ий об'єкт;

$f_k(x,y)$ – залежність частки вилученої інформації на *k*-му об'єкті від співвідношення *x* та *y*.

На першому кроці розглянемо дії нападу. При цьому оптимізаційна задача формулюється так:

$$i(x, y) \to \max, \quad (2)$$

де $x \geq 0,\ y \geq 0,\ \sum_{k=1}^{l} x_k = x,\ \sum_{k=1}^{l} y_k = y.$

В подальшому *x* і *y* позначають сумарні ресурси нападу і захисту, а в функціональних залежностях – незалежні змінні.

Маючи на меті геометричну інтерпретацію розв'язку, розглянемо систему з двох об'єктів. Враховуючи, що складові цільової функції визначаються відношенням $x/y$, введемо нову змінну $\tilde{x} = x/y$. Цільова функція при цьому приймає вигляд:

$$i(\tilde{x}) = \sum_{k=1}^{l} g_k \cdot p_k(\tilde{x}) \cdot q_k(\tilde{x}) \cdot f_k(\tilde{x})$$

де $\tilde{x} = x/y$. В нашому розгляді *x* і *y* перестають бути незалежними змінними і змінюються під дією протилежної сторони. Оберемо залежності $f(\tilde{x})$ у вигляді [4]: $f(\tilde{x}) = \frac{\tilde{x}^n}{\tilde{x}^n + c}$, де параметри *n* і *c* визначають положення і крутизну кривої на різних ділянках. Зокрема, при *n*=1 опуклість кривої $f(\tilde{x})$ направлена догори, при *n*>1 – донизу. При збільшенні *c* крива опускається вниз, причому вплив цього параметру проявляється в більшій степені в початковій області – при $\tilde{x} \lesssim 1$. Прагнучи відобразити особливості залежностей $f(\tilde{x})$, розглянемо функції з різними значеннями *n* і *c*:

$$f(\tilde{x}) = \frac{\tilde{x}}{\tilde{x} + 4} \quad (3)$$

$$f(\tilde{x}) = \frac{\tilde{x}^2}{\tilde{x}^2 + 16} \quad (4)$$

$$f(\tilde{x}) = \frac{\tilde{x}^3}{\tilde{x}^3 + 32} \quad (5)$$

$$f(\tilde{x}) = \frac{\tilde{x}^4}{\tilde{x}^4 + 64} \qquad (6)$$

В подальших розрахунках будемо обирати залежності $f(\tilde{x})$ для двох об'єктів у вигляді пар з приведеного набору (3)-(6). Розглянемо спрощений варіант, в якому $q(\tilde{x}) = const$. З нормовочної умови $\int_0^{\tilde{x}_{гр}} q(\tilde{x})d\tilde{x} = 1$, де $x_{гр}$ обмежує інтервал можливих значень $\tilde{x}$, при обраному $x_{гр} = 3$ одержуємо $q = \frac{1}{3}$.

Розглянемо два варіанта вибору залежностей $f_k(\tilde{x})$ для системи з двох об'єктів: функції (3), (5) (рис. 7.10) і функції (4), (6) (рис. 7.11).

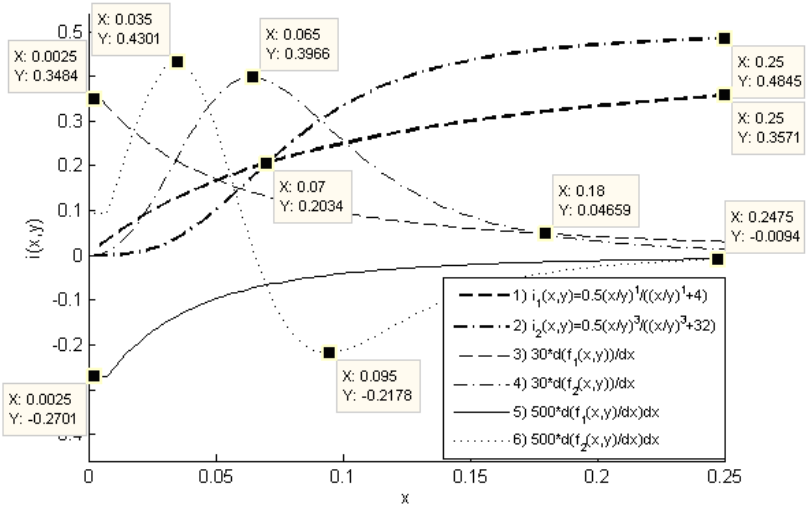



Рис. 7.10.Залежності $i(\tilde{x})$ та похідні від них, одержані на основі дрібно-лінійної (3) і дрібно-нелінійної (5) функцій $f(\tilde{x})$

Покладемо також $g_1 = g_2 = 0{,}5$, $p_1(\tilde{x}) = p_2(\tilde{x}) = 1$ і одержимо:

1) для першого варіанту

$$i(\tilde{x}) = i_1(\tilde{x}) + i_2(\tilde{x}) = \frac{1}{2}\frac{1}{3}\left(\frac{\tilde{x}}{\tilde{x}+4} + \frac{\tilde{x}^3}{\tilde{x}^3+32}\right) \quad (7)$$

2) і для другого

$$i(\tilde{x}) = i_1(\tilde{x}) + i_2(\tilde{x}) = \frac{1}{2}\frac{1}{3}\left(\frac{\tilde{x}}{\tilde{x}+16} + \frac{\tilde{x}^4}{\tilde{x}^4+64}\right). \quad (8)$$

Ці залежності, а також перші та другі похідні залежностей $f(\tilde{x})$, приведені з масштабними коефіцієнтами 30 і , відповідно, 500, зображені на рис. 7.11. Квадратиками позначені екстремальні значення функцій.

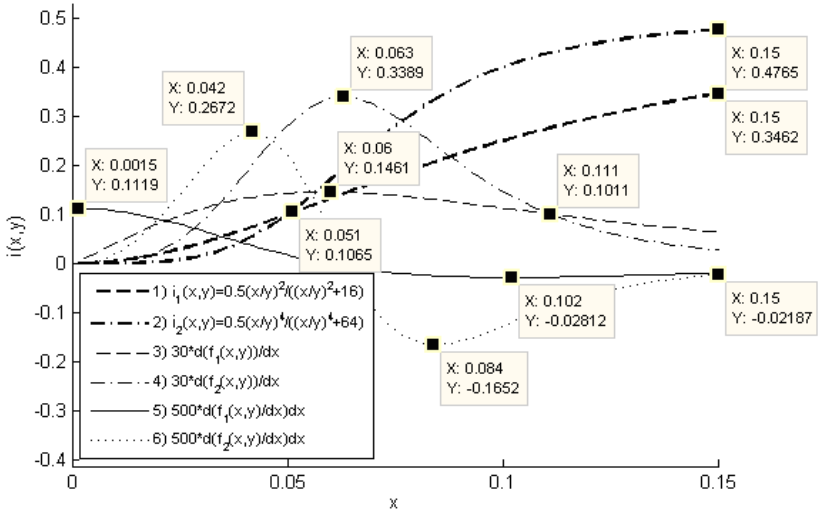



Рис. 7.11.Залежності $i(\tilde{x})$ та похідні від них, одержані на основі дрібно-нелінійних функцій (4) і (6)

Цільова функція $i(x_1, x_2)$ в геометричній інтерпретації зображується у вигляді просторової фігури в системі координат $x_1$, $x_2$, де $x_1$ і $x_2$ – ресурси нападу, направлені на кожний з двох об'єктів (ресурси захисту $y_1$ і $y_2$ вважаються відомими і при розрахунках виступають в ролі параметрів). Фігура побудована на

функціях $i_1(x, y)$ і $i_2(x, y)$ (рис. 7.12). За браком статистичних даних величини, які стоять в правій частині (1) задаються евристично або визначаються шляхом експертної оцінки. При побудові фігури (рис.3) задані такі значення параметрів: $g_1 = g_2 = 0{,}5,\ y_1 = y_2 = 0{,}025$.

Аналізуючи дії кожної з сторін, будемо розрізняти два типи задач – пряму і зворотну. В першому випадку задають кількість ресурсів (нападу і, відповідно, захисту) і визначають їх розподіл по об'єктах, який забезпечує досягнення оптимального значення цільової функції (максимального і, відповідно, мінімального). При рішенні зворотної задачі задається значення цільової функції і потрібно знайти оптимальні значення необхідних ресурсів і їх розподіл по об'єктах.

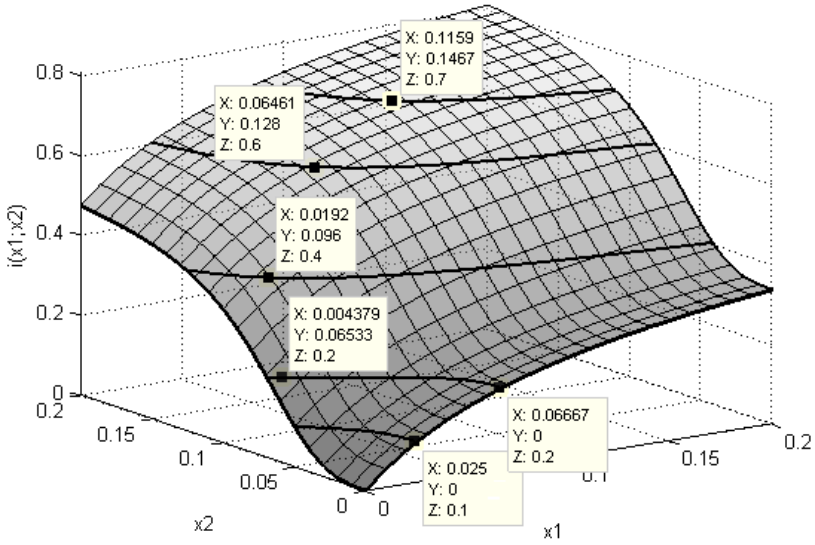



Рис. 7.12. Залежність частки вилученої інформації від ресурсів нападу на двох об'єктах

Геометричний розв'язок прямої задачі одержуємо в результаті перерізу просторової фігури $i(x_1, x_2)$ вертикальною площиною, яка проходить через обмежувальну пряму $x = x_1 + x_2 = C$, розв'язок зворотної – в результаті перерізу цієї фігури

горизонтальною площиною, розташованою на рівні заданого значення *i* (рис. 7.12). В першому випадку оптимум знаходиться в найвищій точці перерізу, в другому – в точці дотику кривої, одержаної в результаті перерізу, до прямої $x_1 + x_2 = x = C$, яка визначає необхідну для досягнення заданого значення *i* кількість ресурсів (рис. 7.13, жирні криві).

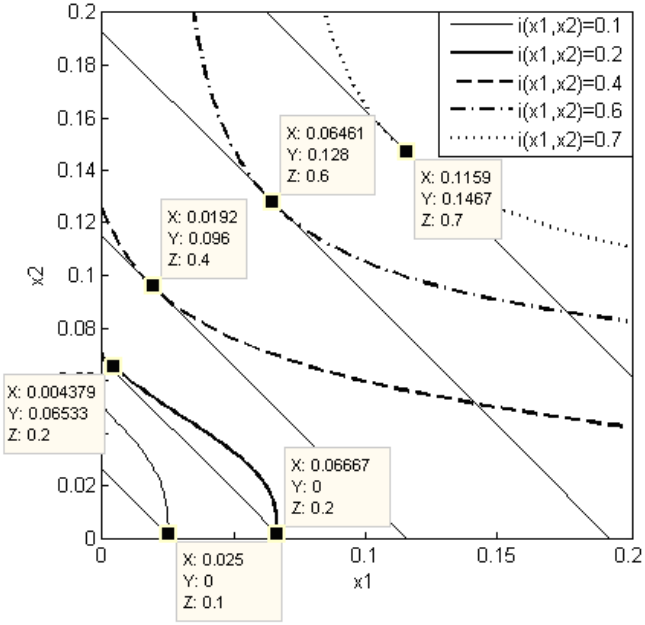



Рис. 7.13.Лінії перерізу просторової фігури (рис. 8.3) на різних рівнях

Положення оптимальних точок (вони зображені квадратиками) дозволяє визначити необхідну кількість ресурсів $x$ – вона визначається діагональною прямою $x = const$, яка є дотичною до кривої перерізу, а відстань між сусідніми прямими характеризує темп зростання $X$ при збільшенні $i(x)$. На цьому ж рисунку показано положення критичної точки, при досягненні якої слід переходити від зосередження ресурсів нападу на одному об'єкті до їх розподілу між обома об'єктами (на рис. 7.13 $x_{кр} = 0.067$, а $i_{кр} = 0.2$).

Лінія, яка з'єднує оптимальні точки, є результат розрахунків зворотної задачі, в якій по заданим значенням $i(x_1, x_2)$ знаходять оптимальні значення $x_1^0, x_2^0, x^0$. Ці лінії зображені на рис. 7.14-16, в яких використані різні залежності $f_k(x, y)$. Розв'язки оптимізаційних задач одержано з допомогою пакету Optimization Toolbox програмного комплексу Matlab.

При аналізі залежностей (рис. 7.14-7.16) нас в першу чергу цікавить положення критичних точок, які відображають зміну стратегії в розподілі ресурсів. Перші дві з них – $x_{кр1}$ і $x_{кр2}$ визначають перехід від концентрації ресурсів на одному з об'єктів до їх зосередження на іншому ($x_{кр1}$) або до розподілу між обома об'єктами ($x_{кр2}$) (рис. 7.15). Третя характерна точка відображає не зміну стратегії, а відображає лише якісну зміну переваги в розподілі ресурсів від одного об'єкта до іншого (для однотипності позначимо її через $x_{кр3}$). На рис. 7.15 $x_{кр1}$ відсутня, $x_{кр2} = 0{,}069$, $x_{кр3} = 0{,}356$; на рис. 6 $x_{кр1} = 0{,}049$, $x_{кр2} = 0{,}136$, $x_{кр3} = 0{,}215$.

Вплив параметрів $n$ і $c$ в залежностях $f_k(x, y)$ на положення критичних точок можна сформулювати наступним чином.

1. При збільшенні $n$ значення $x_{кр i}$ ($i = 1,2,3$) зменшуються, $i_{кр1}$ та $i_{кр3}$ зменшуються, а $i_{кр2}$ зростають (рис. 7.13 і 7.14).

2. При збільшенні $c$ і $n_1 \neq n_2$ $x_{кр i}$ зростають, $i_{кр2}$ також зростає, а $i_{кр3}$ зменшується.

Отже, при збільшенні $n$ і збільшенні $c$ значення $x_{кр i}$ змінюються різнонаправлено.

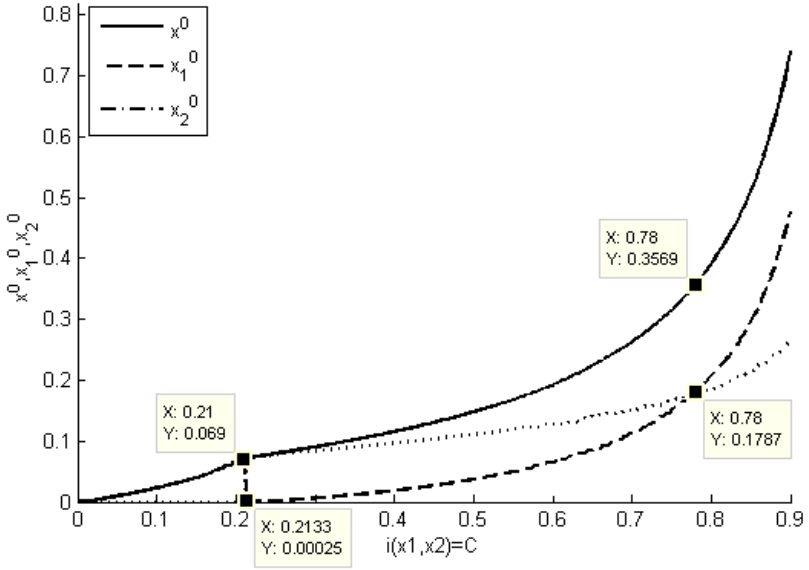



Рис. 7.14. Оптимальний розподіл ресурсів нападу між двома об'єктами, які характеризуються функціями (3), (5)

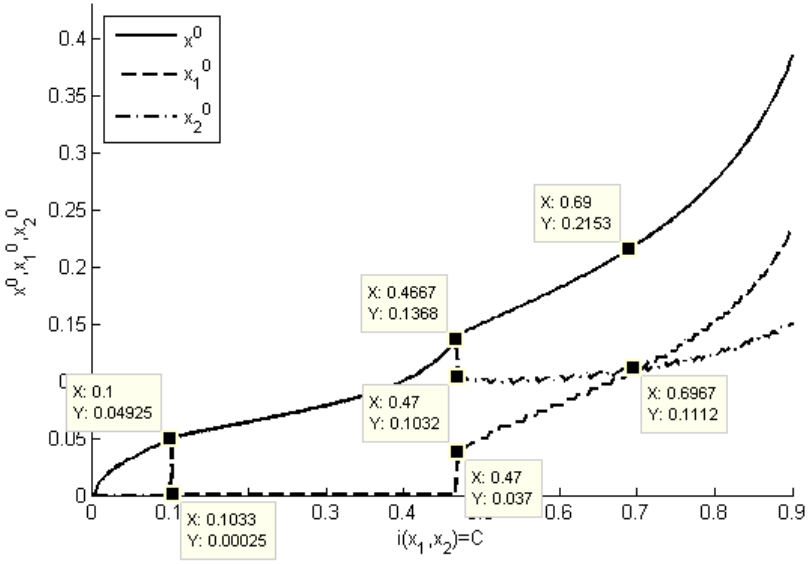



Рис. 7.15.Оптимальний розподіл ресурсів нападу між двома об'єктами, які характеризуються функціями (4), (6)

Положення критичних точок визначається кривизною функцій $f_1(\tilde{x})$ і $f_2(\tilde{x})$ та відображає перехід до нового принципу розподілу, який забезпечує більший приріст значень цільової функції при подальшому зростання $x$. Зокрема, при $x = x_{кр1}$ $f_1(x_{кр1}) + \frac{df_1}{dx_1} dx_1 = f_2(x_{кр1}) + \frac{df_2}{dx_2} dx_2$ і при подальшому зростанні $x$ права частина рівності перевищує ліву. При переході через другу критичну точку $x_{кр2}$ значення $x$ можна поділити на дві складові $x_1$ і $x_2$, $x_1 + x_2 = x$ – такі, що сума $f_1(x_1) + \frac{df_1}{dx_1} dx_1 + f_2(x_2) + \frac{df_2}{dx_2} dx_2$ стає більшою від кожної з величин $f_1(x) + \frac{df_1}{dx} dx$ і $f_2(x) + \frac{df_2}{dx} dx$.

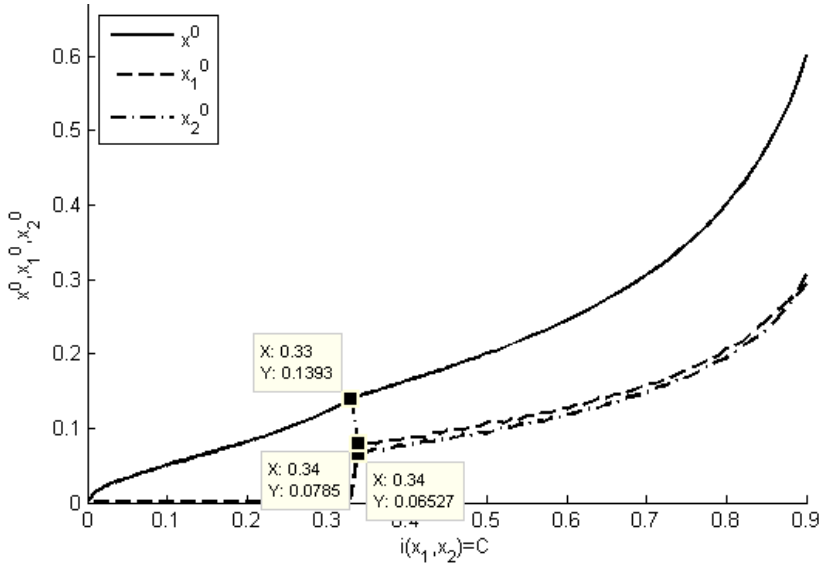



Рис. 7.15.Оптимальний розподіл ресурсів нападу між двома об’єктами з однаковими залежностями $f_k(x)$ (4)

На рис. 7.13 перша критична точка $x_{кр1}$ відсутня. Це характерно для систем, в яких один з об'єктів описується дробно-лінійною функцією $f(\tilde{x}) = \frac{\tilde{x}}{\tilde{x} + c}$. Відзначимо також, що подібна ситуація спостерігається у випадку двох однакових об'єктів: існує одна критична точка $x_{кр}$, причому при $x < x_{кр}$ ресурси слід концентрувати на одному з об'єктів, а при $x > x_{кр}$ – поділяти порівну між об'єктами (рис. 7.15).

Аналізуючи положення характерних точок, можна зробити такі висновки.

1. Точка $x_{кр1}$ близька до точки перетину кривих $f_1(x_1)$, $f_2(x_2)$ ($x_{кр1} = 0{,}069$ і $x_{кр1} = 0{,}049$, точка перетину кривих $x = 0{,}070$ і, відповідно, $x = 0{,}051$).

2. Точка $x_{кр3}$ близька до точки перетину похідних $f_1'(x_1)$, $f_2'(x_2)$ ($x_{кр3} = 0{,}179$ і $x_{кр3} = 0{,}111$, точки перетину похідних $x = 0{,}180$ і, відповідно, $x = 0{,}111$).

3. Значення $x_{кр2}$ дещо перевищує $x_{кр3}$, а після «стрибка» $x_2$ стає близьким до нього ($x_2 = 0{,}103$, а $x_{кр3} = 0{,}111$).

Приведені результати дозволяють оцінити кількість і розподіл ресурсів нападу, що буде корисним при розробці ефективних заходів протидії. Використовуючи описану методику і знайдені ресурси нападу, можемо визначити оптимальний розподіл ресурсів захисту. Продовжуючи цю процедуру, зрештою прийдемо до динамічного управління ресурсів в інформаційному протистоянні.

## 7.4 Аналітичні методи

### *7.4.1 Перехід від геометричного метода до алгебраїчного.*

При графічній інтерпретації задач лінійного програмування зазначалось, що оптимальний розв'язок досягається в одній з

кутових точок простору рішень. Це є ключовою ідеєю при розробці алгебраїчного метода розв'язку таких задач. Перехід від графічного до алгебраїчного методу (останній отримав назву симплекс-метода) здійснюється через алгебраїчне представлення кутових точок. Розглянемо сутність методу на прикладі.

*Приклад.* Знайти $z=2x_1+3x_2 \rightarrow max$ при обмеженнях

$$\begin{cases} 2x_1 + x_2 \le 4 \\ x_1 + 2x_2 \le 5 \\ x_1 \ge 0,\ \ x_2 \ge 0 \end{cases} \qquad (1)$$

Коефіцієнти при змінних $x_1$ і $x_2$ у виразі для цільової функції $Z(x_1,x_2)$ визначають кількість інформації, вилученої з першого та, відповідно, другого об'єкта, на одиницю витрат.

Перейдемо від нерівностей у формулюванні обмежень до рівнянь. Для цього введемо нові, додаткові змінні $x_3$ і $x_4$, котрі й будуть забезпечувати рівність лівих і правих частин:

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 \\ x_1 + 2x_2 + x_4 = 5 \\ x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0,\ x_4 \ge 0 \end{cases} \qquad (2)$$

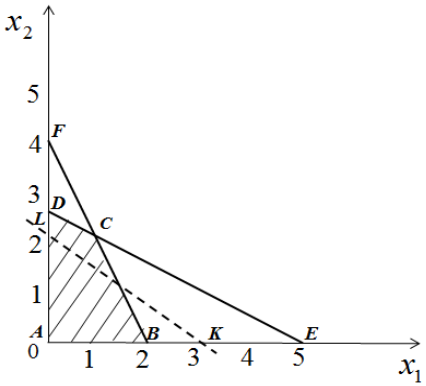



Рис.7.16 Геометрична інтерпретація симплекс-метода. Простір розв'язків для приведеного прикладу

Маємо систему двох рівнянь з чотирма змінними. Очевидно, двом з цих чотирьох змінних можна надати певні значення і одержати систему двох рівнянь з двома змінними. Якщо одній зі змінних надається нульове значення, то ми одержуємо рівняння однієї з границь області рішень, якщо нульові значення надаються обом змінним, то цим визначається кутова точка. Наприклад, при $x_1=0$ в системі координат

$x_1$,$x_2$одержуємо з (1) $x_2=4-x_3$ (або $x_2=\frac{5-x_4}{2}$) – це рівняння осі ординат, котра є однією з границь області *ABCD*. При $x_1=0$,$x_2=0$ одержуємо кутову точку *A*, при $x_1=0$,$x_3=0$ – точку *F*, в якій $x_2=4$, проте це неприпустима кутова точка, оскільки в ній не виконується четверта з нерівностей (2): в цій точці, як випливає з другого рівняння (1) $x_4=-3$. Змінні, яким надаються нульові значення при знаходженні кутових точок, називаються небазисними, а інші змінні - базисними. Якщо всі базисні змінні приймають невід'ємні значення, то вони формують припустиме базисне рішення, в протилежному випадку – неприпустиме. В математичній літературі замість термінів «базисне рішення» і «припустиме базисне рішення» вживають також терміни «план» і «опорний план».

В табл.4 приведені результати рішення поставленої задачі.

Таблиця 7.1

Результати рішення задачі

| Небазисні змінні | Базисні змінні | Базисні розв'язки | Відповідні кутові точки | Припустимі кутові точки | Значення цільової функції *z* |
|---|---|---|---|---|---|
| $(x_1, x_2)$ | $(x_3, x_4)$ | (5, 4) | *A* | *A* | 0 |
| $(x_1, x_3)$ | $(x_2, x_4)$ | (4, -3) | *F* | – | – |
| $(x_1, x_4)$ | $(x_2, x_3)$ | (2.5, 1.5) | *D* | *D* | 7.5 |
| $(x_2, x_3)$ | $(x_1, x_4)$ | (2, 3) | *B* | *B* | 4 |
| $(x_2, x_4)$ | $(x_1, x_3)$ | (5, -6) | *E* | – | – |
| $(x_3, x_4)$ | $(x_1, x_2)$ | (1, 2) | *C* | *C* | 8 |

Як видно з табл.7.1, розв'язок $z_{max}=8$ знаходиться в точці *C (1, 2)*. Цей результат можна також одержати з графічного рішення, переміщуючи лінію рівня *KL* в бік віддалення від початку координат до досягнення крайньої точки області допустимих значень (рис.7.16).

Таким чином, геометричний зміст симплекс - метода полягає в послідовному переході від однієї вершини многокутника обмежень до сусідньої, в котрій цільова функція приймає краще значення – до того моменту, поки не буде знайдено оптимальне рішення в вершині, де досягається оптимальне значення цільової функції.

При застосуванні симплекс-методу ми стоїмо перед необхідністю рішення системи *m*рівнянь з $n>m$ невідомими.

Розв'язок знаходиться в результаті почергового прирівнювання $n-m$ змінних до нуля, чим визначаються окремі кутові точки. В кожній з цих точок ми знаходимо значення цільової функції і, порівнюючи ці значення, визначаємо оптимальний розв'язок. Кількість комбінацій, які необхідно розглянути при простому перебиранні змінних, становить $C_n^m = \frac{n!}{m!(n-m)!}$ і дорівнює кількості кутових точок. В нашій задачі ми мали $C_4^2 = \frac{4!}{2!2!} = 6$ кутових точок і відповідних рівнянь. При $m=4$, $n=8$ матимемо $C_8^4 = \frac{8!}{4!4!} = 70$ рівнянь, при $m=10$, $n=20$ (а в реальних умовах зустрічаються і більші значення $min$) – 184756 рівнянь. Розв'язок такої системи являє собою вражаючу в обчислювальному сенсі задачу. Симплекс-метод направлений на спрощення цієї процедури за рахунок того, що при перебиранні можливих варіантів ми відкидаємо неприпустимі базисні рішення і розглядаємо лише частину всіх допустимих. Це досягається в результаті використання оптимізованої ітераційної процедури і побудові відповідного алгоритму.

### *7.4.2. Обчислювальна схема симплекс-метода*

Проілюструємо застосування ітераційної методики на прикладі, розглянутому в 7.4.1. Цільову функцію $z$ запишемо у вигляді рівняння

$$z - 2x_1 - 3x_2 = 0, \tag{4}$$

до якого додамо систему рівнянь, що випливають з обмежень

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 + 2x_2 + x_4 &= 5, \end{aligned} \tag{5}$$

і умови невід'ємності змінних

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \tag{6}$$

Коефіцієнти при змінних у рівняннях (4), (5) формують початкову розрахункову таблицю 1. В якості базисних змінних

вибираємо змінні, кожна з яких тільки один раз входить у рівняння. В нашому прикладі це $x_3$, $x_4$. Решті змінних надаємо нульові значення:$x_1$=0, $x_2$=0. Це є базисні, або нульові змінні. Умови невід‘ємності (6) відносимо до неосновних обмежень. Вони в таблицях не відображені. Номер рядка – це номер базисної змінної в даній таблиці. Нульовий рядок відповідає цільовій функції (4) і служить для визначення ступеня оптимальності припустимого базисного рішення.

Таблиця7.2

| Номер стовпчика | | 0 | | | | |
|---|---|---|---|---|---|---|
| Номер рядка | Базисні змінні | Базисне рішення | $x_1$ | $x_2\downarrow$ | $x_3$ | $x_4$ |
| 0 | $z$ | 0 | -2 | -3 | 0 | 0 |
| 1 | $x_3$ | 4 | 2 | 1 | 1 | 0 |
| 2 | $x_4\rightarrow$ | 5 | 1 | 2 | 0 | 1 |

На першому кроці ми поклали $x_1$=0, $x_2$=0 − це точка $A(0;0)$ на рис. 7.10. Значення $x_3$=4, $x_4$=5, знайдені з (5) і відображені в стовпчику «базисне рішення», не впливають на положення точки в координатах$x_1$, $x_2$.

Розподіл ресурсів $\{x_k\}, k=\overline{1,4}$ вважається оптимальним, якщо в задачі максимізації в нульовому рядку серед коефіцієнтів при змінних відсутні від‘ємні числа (в задачі мінімізації ця умова замінюється на умову відсутності додатних чисел). В табл.1 в нульовому рядку маємо два від’ємних числа: (-2) при змінній $x_1$ і (-3) при змінній $x_2$. Отже, варіант, відображений в табл. 7.2, не є оптимальним, і точка $A$ (0;0) не є точкою максимуму. Необхідно продовжити пошук.

Наша наступна задача – вибрати найкоротший шлях з початкової точки $A(0;0)$ до точки максимуму. Можливі два напрямки руху вздовж периметра многокутника рішень – за стрілкою годинника і проти її руху. На першому кроці ці напрямки визначають: у першому випадку рух вздовж осі $x_2$, а в другому – вздовж осі $x_1$. Оскільки за умовою необхідно досягти максимуму цільової функції, будемо рухатись з точки $A(0;0)$ в напрямку найбільшого зростання функції $z$. Швидкість зростання функції

$z = 2x_1 + 3x_2$ при збільшенні $x_1$ і $x_2$ визначається коефіцієнтами при цих змінних. Більшим є коефіцієнт при $x_2$: він дорівнює 3. Отже, будемо рухатись з точки $A$ вздовж осі $x_2$. При цьому треба визначити, в якій точці необхідно зупинитись. Це одна з граничних точок, які визначаються обмежувальними нерівностями (1). Їх положення знайдемо з рівнянь границь, які одержуємо, переходячи від нерівностей до рівнянь

$$2x_1 + x_2 = 4, \tag{7}$$

$$x_1 + 2x_2 = 5, \tag{8}$$

Враховуючи, що $x_1$=0, маємо: з (7) $x_2^{(1)} = 4$ (точка$E$), з (8) $x_2^{(2)} = \frac{5}{2}$(точка $B$). Звертаючись до табл. 7.2, ці значення можна отримати, знаходячи відношення чисел стовпчика «базисне рішення» до відношення додатних чисел ключового стовпчика, в якому відображені значення $x_2$: $4 \div 1 = 4$, $5 \div 2 = \frac{5}{2}$. Оскільки нерівностями (2), (3) ми обмежені зверху, то з двох можливих точок $x_2^{(1)}$, $x_2^{(2)}$ обираємо точку з меншим значенням $x_2^{(2)} = \frac{5}{2}$. Це точка $B$. Елемент таблиці, який відповідає цьому значенню, (це коефіцієнт при $x_2$ в рівнянні (8)) є ключовим. В табл. 6.2 це число 2 з другого рядка (воно взяте в рамочку).

Таким чином, змінна $x_1$ залишається вільною ($x_1$=0), а змінна $x_2$ переходить в розряд базисних. Виникає питання: яку з базисних змінних замінить $x_2$, $x_3$ чи$x_4$. Це визначає рядок, в котрому знаходиться ключовий елемент – це$x_4$ .

Сформуємо другу розрахункову таблицю, яка відповідає точці $B$ і використовується при подальших розрахунках. Для цього виконаємо такі операції:

1. Замість базової змінної $x_4$ ключового рядка вводимо нову базову змінну $x_2$ ключового стовпчика.

2. Ключовий рядок одержуємо від ділення його елементів попередньої таблиці на ключовий елемент, тобто на 2 : $\frac{5}{2} = 2{,}5$; $\frac{1}{2} = 0{,}5$; $\frac{2}{2} = 1$; $\frac{0}{2} = 0$; $\frac{1}{2} = 0{,}5$.

3. Решту комірок заповнюємо за правилом прямокутника :

$$a'_{ij} = \frac{a_{ij}a_{qs} - a_{is}a_{qj}}{a_{qs}},$$

де $a'_{ij}$ − шуканий елемент нової таблиці, $a_{ij}$ − попередньої, $a_{qs}$ − ключовий елемент, $i, q$ − номери рядків ( від 0 до 2), $j, s$ − номери стовпчиків (від 0 до 4). Оскільки ключовий елемент знаходиться у другому рядку і другому стовпчику,то $q = 2$, $s = 2$ і$a_{qs} = a_{22} = 2$.

Наведемо результати розрахунків.

Нульовий рядок:

$$a'_{00} = \frac{a_{00}a_{22} - a_{02}a_{20}}{a_{22}} = \frac{0 \cdot 2 - (-3) \cdot 5}{2} = 7{,}5$$

$$a'_{01} = \frac{a_{01}a_{22} - a_{02}a_{21}}{a_{22}} = \frac{(-2) \cdot 2 - (-3) \cdot 1}{2} = -0{,}5$$

$$a'_{02} = \frac{a_{02}a_{22} - a_{02}a_{22}}{a_{22}} = 0$$

$$a'_{03} = \frac{a_{03}a_{22} - a_{02}a_{23}}{a_{22}} = \frac{0 \cdot 2 - (-3) \cdot 0}{2} = 0$$

$$a'_{04} = \frac{a_{04}a_{22} - a_{02}a_{24}}{a_{22}} = \frac{0 \cdot 2 - (-3) \cdot 1}{2} = -1{,}5$$

Перший рядок:

$$a'_{10} = \frac{a_{10}a_{22} - a_{12}a_{20}}{a_{22}} = \frac{4 \cdot 2 - 1 \cdot 5}{2} = 1{,}5$$

$$a'_{11} = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{22}} = \frac{2 \cdot 2 - 1 \cdot 1}{2} = 1{,}5$$

$$a'_{12} = \frac{a_{12}a_{22} - a_{12}a_{22}}{a_{22}} = \frac{1 \cdot 2 - 1 \cdot 2}{2} = 0$$

$$a'_{13} = \frac{a_{13}a_{22} - a_{12}a_{23}}{a_{22}} = \frac{1 \cdot 2 - 1 \cdot 0}{2} = 1$$

$$a'_{14} = \frac{a_{14}a_{22} - a_{12}a_{24}}{a_{22}} = \frac{0 \cdot 2 - 1 \cdot 1}{2} = -0{,}5$$

Заповнюємо табл. 7.3 і перевіряємо одержане базисне рівняння на оптимальність. В нульовому рядку є від'ємне число (-0,5) у стовпчику$x_1$. Це ключовий стовпчик. Беремо відношення чисел базисного рішення для$x_3$,$x_2$ до відповідних чисел ключового стовпчика:

1) 1,5:1,5=1
2) 2,5:0,5=5

Меншим з цих чисел є перше число. Отже, ключовим елементом є число 1,5 ключового стовпчика. Нова базисна змінна $x_1$, яка визначається ключовим стовпчиком, замінює змінну $x_3$, яка визначається ключовим елементом.

Таблиця 7.3

| Номер стовпчика | | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| Номер рядка | Базисні змінні | Базисне рішення | $x_1 \downarrow$ | $x_2$ | $x_3$ | $x_4$ |
| 0 | $z$ | 7,5 | -0,5 | 0 | 0 | 1,5 |
| 1 | $x_3 \rightarrow$ | 1,5 | 1,5 | 0 | 1 | -0,5 |
| 2 | $X_2$ | 2,5 | 0,5 | 1 | 0 | 0,5 |

В геометричній інтерпретації необхідність продовження пошуку пояснюється тим, що точка, яка відображається даними табл. 7.3 (це точка $B$ з координатами $x_1$=0, $x_2$=2,5) не відповідає оптимальному рішенню. Значення $z$=7,5 , як ми побачимо, не є оптимальним. Використовуючи наведену методику, виконуємо необхідні розрахунки і заповнюємо табл. 7.4.

Таблиця 7.4

| Номер стовпчика | | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| Номер рядка | Базисні змінні | Базисне рішення | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
| 0 | $z$ | 8 | 0 | 0 | $1/3$ | $1\,1/3$ |
| 1 | $X_1$ | 1 | 1 | 0 | $2/3$ | $-1/3$ |
| 2 | $X_2$ | 2 | 0 | 1 | $-1/3$ | $2/3$ |

В нульовому рядку вже відсутні від'ємні числа. Отже, одержане базисне рішення є оптимальним. В геометричному плані воно зображується точкою С (рис. 7.16) з координатами $x_1$=1, $x_2$=2. Економічний смисл розв'язку полягає в тому, що він визначає оптимальний розподіл ресурсів $x_1^0$: $x_2^0$=1:2 при заданих обмеженнях (1) на ресурси. Максимальне значення вилученої інформації становить при цьому $z_{max}$=8.

### *7.4.3 Метод Лагранжа*

В попередніх параграфах розглянуті методи оптимізації лінійних задач. В реальних умовах цільова функція має дрібно-лінійний або дрібно-нелінійний характер. Розглядаючи, як і раніше, систему, котра містить два об'єкти, проілюструємо можливість аналітичних методів. При дрібно-лінійній формі цільова функція для сторони нападу у найпростішому випадку має такий вигляд:

$$i(x_1, x_2) = g_1 \frac{x_1 / y_1}{x_1 / y_1 + c_1} + g_2 \frac{x_2 / y_2}{x_2 / y_2 + c_2},$$

де $y_1$, $y_2$виступають в ролі параметрів.

Перший з аналітичних методів – метод Якобі – дозволяє, використовуючи обмежувальну рівність $x_1 + x_2 = X$, звести задачу на умовний екстремум функції двох змінних до задачі на безумовний екстремум функції однієї змінної:

$$i(x_1) = g_1 \frac{x_1 / y_1}{x_1 / y_1 + c_1} + g_2 \frac{(X - x_1) / y_2}{(X - x_1) / y_2 + c_2},$$

Умовою оптимальності є рівність

$$\frac{di(x_1)}{dx_1} = g_1 \frac{x_1 / y_1}{(x_1 / y_1 + c_1)^2} + g_2 \frac{(X - x_1) / y_2}{\left[(X - x_1) / y_2 + c_2\right]^2} = 0$$

Після елементарних перетворень одержуємо досить громіздке квадратне рівняння, з котрого визначаємо оптимальне значення $x_1^0$, а потім і $x_2^0$. При дрібно-нелінійній формі функції *i(x,y)* маємо алгебраїчне рівняння більш високого ступеня, рішення котрого викликає значні обчислювальні труднощі.

Звертаючись до другого з аналітичних методів – методу Лагранжа, - будуємо функцію

$$L(x_1, x_2, \lambda) = g_1 \frac{x_1 / y_1}{x_1 / y_1 + c_1} + g_2 \frac{x_2 / y_2}{x_2 / y_2 + c_2} + \lambda(x_1 + x_2 - X),$$

де введена третя змінна – невідомий множник Лагранжа $\lambda$.

Прирівнюємо перші похідні по змінним до нуля:

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = g_1 \dfrac{c_1 / y_1}{(x_1 / y_1 + c_1)^2} + \lambda = 0 \\ \dfrac{\partial L}{\partial x_2} = g_2 \dfrac{c_2 / y_2}{(x_2 / y_2 + c_2)^2} + \lambda = 0 \\ \dfrac{\partial L}{\partial \lambda} = x_1 + x_2 - X = 0 \end{cases}$$

Звідси одержуємо вирази для оптимальних значень $x_1^0$, $x_2^0$ і значення $\lambda$:

$$x_1^0 = \sqrt{\frac{c_1 g_1 y_1}{\lambda}} - c_1 y_1$$

$$x_2^0 = \sqrt{\frac{c_2 g_2 y_2}{\lambda}} - c_2 y_2$$

$$\lambda = \frac{(\sqrt{g_1 y_1} + \sqrt{g_2 y_2})^2}{(X + c_1 y_1 + c_2 y_2)^2}$$

Наведені вирази легко розповсюдити на довільну кількість об'єктів.

При розгляді дій захисту метод Лагранжа приводить до аналогічних виразів, які в загальному випадку мають такий вигляд:

$$y_k^0 = \sqrt{\frac{c_k g_k y_k}{\lambda}} - c_k x_k, \text{ де } \lambda = \frac{(\sum_{k=1}^{l} c_k g_k x_k)^2}{(X + Y)^2}.$$

Наведемо приклад застосування методу Лагранжа.

Нехай $c_1$=$c_2$=1, $y_1$=$y_2$=1, $X$=1. Знайти оптимальний розподіл ресурсів нападу в залежності від розподілу інформації на об'єктах.

I варіант

$g_1$=$g_2$=0,5

Підставивши ці величини у вирази для оптимальних значень $x_1^0$, $x_2^0$ і значення $\lambda$, отримаємо:

$x_1^0 = x_2^0 = 0{,}5$;

$\lambda$=0,22

II варіант

$g_1$=0,4, $g_2$=0,6

$x_1^0 = 0{,}35$; $x_2^0 = 0{,}65$;

$\lambda$=0,22

III варіант

$g_1$=0,2, $g_2$=0,8

$x_1^0 = 0;\ x_2^0 = 1;$

$\lambda$=0,2

Як бачимо, у перших двох варіантах ресурси захисту слід розподіляти між двома об'єктами, а в третьому – зосередити на другому об'єкті.

### Контрольні питання

1. Умови досягнення екстремуму для різних видів функцій.
2. Сідлова точка, її сутність і роль в задачах інформаційної безпеки.
3. Зображення екстремумів цільових функцій нападу і захисту.
4. Сутність геометричного методу рішення задач лінійного програмування.
5. Форма многокутника допустимих розв'язків для нападу і захисту.
6. Принципи визначення чутливості рішення до зміни вхідних даних при його геометричній інтерпретації.
7. Симплекс-метод, його обчислювальна схема і застосування в задачах інформаційної безпеки.
8. Метод Лагранжа, його сутність і застосування.
9. Закономірності оптимального розподілу ресурсів нападу між двома об'єктами при різних варіантах функцій *f(x)*.

## VIII. ЕФЕКТИВНІСТЬ РОЗВІДКИ ПРИ ПРОТИСТОЯННІ ДВОХ СТОРІН В ІНФОРМАЦІЙНІЙ СФЕРІ

### 8.1 Постановка задачі

Питання про ефективність розвідки розглядалось в [5], проте постановка задачі і прийняті допущення значно ускладнюють використання одержаних результатів (зокрема, це стосується прийнятої моделі Гроса).

Ми використаємо модель [13]. Цільову функцію, яка визначає кількість вилученої інформації, представимо у вигляді:

$$i(\tilde{x}) = \sum_{k=1}^{l} g_k \cdot p_k \cdot q_k(\tilde{x}) \cdot f_k(\tilde{x}), \qquad (8.1)$$

де $\tilde{x} = x/y$ – відносна величина, яка характеризує співвідношення ресурсів нападу і захисту – $x$ і ,відповідно, $y$;

$g_k$ – відносна кількість інформації на $k$-му об'єкті;

$p_k$ – імовірність нападу на $k$-й об'єкт;

$q_k(\tilde{x})$ – імовірність виділення нападом ресурсів $x$ на $k$-ий об'єкт;

$f_k(\tilde{x})$ – залежність частки вилученої інформації від ресурсів $x$ та $y$.

Розглянемо спрощений варіант, коли система складається з двох об'єктів, причому $p_k = 1;\ g_1 = g_2 = g/2 = \frac{1}{2};\ q_k(\tilde{x}) = q = 1/3$.

Останнє значення знаходимо з умови $\int_0^{\tilde{x}_{гр}} q(\tilde{x})d\tilde{x} = 1,$ де $\tilde{x}_{гр} = 3$ – границя інтервалу можливих, на наш погляд, значень $\tilde{x}$.

Відносна кількість вилученої інформації з двох об'єктів становить:

$$i(\tilde{x}) = \frac{1}{2} \cdot \frac{1}{3}\left(f_1(\tilde{x}) + f_2(\tilde{x})\right). \qquad (8.2)$$

Нагадаємо, що залежності $f(x)$ можна описати двома типами функцій – степеневою $f(\tilde{x}) = \frac{a\tilde{x}^n}{\tilde{x}^n + c}$ і показниковою $f(\tilde{x}) = 1 - e^{-m\tilde{x}^n}$, де сталі $a, c, n, m$ визначають положення і нахил кривих. Враховуючи, що при певному виборі параметрів ці залежності можуть стати досить близькими, обмежимось в подальшому розгляді функціями першого типу.

## 8.2 Результати досліджень

На рис. 8.1-8.10 приведені результати розрахунків, виконаних з використанням пакету Optimization Toolbox програмного комплексу Matlab при різних видах функцій $f(\tilde{x})$ та різних значеннях кількості ресурсів нападу $X = X^{(1)} + X^{(2)}$, де $X^{(1)}$ і $X^{(2)}$ – ресурси, направлені на розвідку і, відповідно, на вилучення інформації. В подальших розрахунках і на рисунках прийнято, що загальний ресурс захисту $Y = 0{,}05$ рівномірно розподілений між об'єктами: $y_1 = y_2 = 0{,}025$. На лівих частинах рисунків приведені залежності $i(\tilde{x})$ та їх похідні для кожного з двох об'єктів, які описуються степеневими функціями $f(\tilde{x}) = \frac{\tilde{x}^n}{\tilde{x}^n + c}$ в різних інтервалах зміни загальної кількості ресурсів $x$ – від 0 до 0,05, що відповідає максимальному відношенню $\frac{x_{\max}}{y} = \frac{0{,}05}{0{,}025} = 2$, і від 0 до 0,2 ($\frac{x_{\max}}{y} = \frac{0{,}2}{0{,}025} = 8$). Підкреслимо, що в області $x \gtrsim 0$ ці функції мають різний характер: при $n = 1$ опуклість функції $f(\tilde{x})$ направлена вгору, а при $n > 1$ – вниз, що відображає різну вразливість об'єктів в початковій області і дозволяє виявити вплив цього фактору. Вибір параметрів $a, c$ в наших розрахунках не має принципового значення і обумовлений бажанням найбільш яскраво відобразити описані нижче закономірності.

На правих частинах рисунків – втрати інформації з обох об'єктів під час розвідки, під час нападу і сумарні. По осі абсцис відкладені ресурси, вкладені в кожний об'єкт під час розвідки. Ми вважаємо, що протистояння здійснюється в умовах повної невизначеності і ресурси, виділені на розвідку, діляться між об'єктами порівну: $x_1^{(1)} = x_2^{(1)} = x^{(1)}$, загальний ресурс розвідки $X^{(1)} = x_1^{(1)} + x_2^{(1)} = 2x^{(1)}$, залишок ресурсів після розвідки використовується на напад. Максимальні значення на осі абсцис

правих рисунків вдвічі менші, ніж на лівих (ці значення відображають той крайній випадок, коли всі ресурси вкладають порівну між об'єктами). Точки $x = 0$ на правих рисунках відповідають іншій граничній ситуації – всі ресурси вкладають в один з об'єктів. Квадратики на кривих – екстремальні значення залежностей.

Наші дослідження направленні на виявлення ролі двох основних факторів, які впливають на ефективність розвідки:

1) вразливості кожного з об'єктів, яка виражається залежністю $f_k(\tilde{x})$;

2) загального розміру $X$ ресурсів нападу.

Кінцевою метою є визначення доцільності проведення розвідки в кожній конкретній ситуації і у випадку позитивного рішення цього питання – визначення оптимального співвідношення $X^{(1)} \big/ X^{(2)}$ ресурсів, виділених на розвідку і на вилучення інформації, а також оптимального розподілу ресурсів між об'єктами. Критерієм оптимуму є досягнення максимальної кількості вилученої інформації $i(\tilde{x})$.

Застосування цього критерію в умовах невизначеності потребує деяких зауважень.

1. Вважаємо, що після проведення розвідки напад робить правильний вибір об'єкта, на котрий слід направляти весь залишок ресурсів, і кількість вилученої інформації визначається на кожній дільниці верхньою з двох кривих, які зображають сумарний витік.

2. Єдиною точкою, де існує повна визначеність відносно кінцевого результату є точка $X^{(1)} = X^{(1)}_{\max}$ (на рис.7.1,б це $x^{(1)} = 0{,}025$), в якій два етапи (розвідка і вилучення) зливаються в один, тобто фактично розвідка не проводиться, а всі ресурси направляються порівну на об'єкти. Значення $i_{\Pi}(\tilde{x})$ в цій точці і буде орієнтиром при вирішенні питання про доцільність проведення розвідки.

З врахуванням приведених міркувань на рисунках вся шкала $x$ поділена на дві дільниці: перша (позначається лівосторонньою штриховкою) – це зона, в якій розвідка доцільна (штрихова жирна

лінія на рис.7.1а лежить вище горизонтальної лінії, проведеної на рівні $x_{\Pi} = 0{,}06667$) і друга, в якій розвідка недоцільна (правостороння штриховка). На рис.8.1,2 зона доцільності розташована в лівій частині інтервалу зміни $x$, на рис.8.4 – всередині, на рис.8.5 – в правій частині, на рис.8.3 вона займає всю шкалу, а на рис.8.6 – взагалі відсутня.

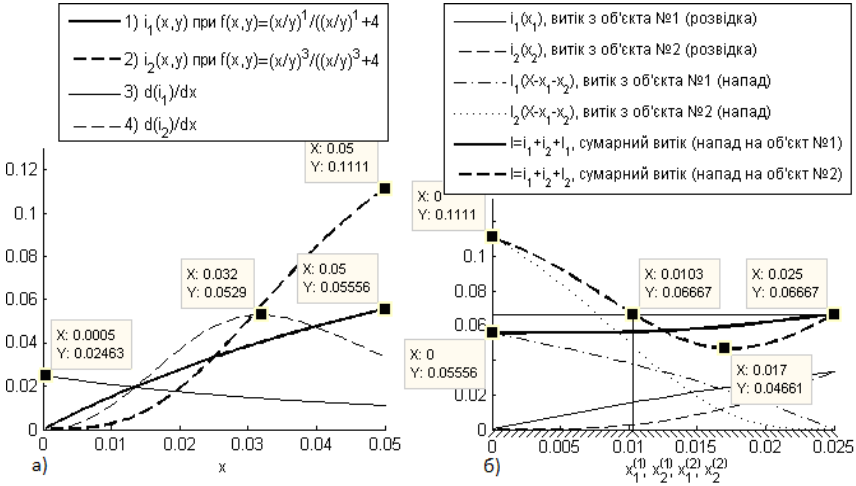



Рис.8.1

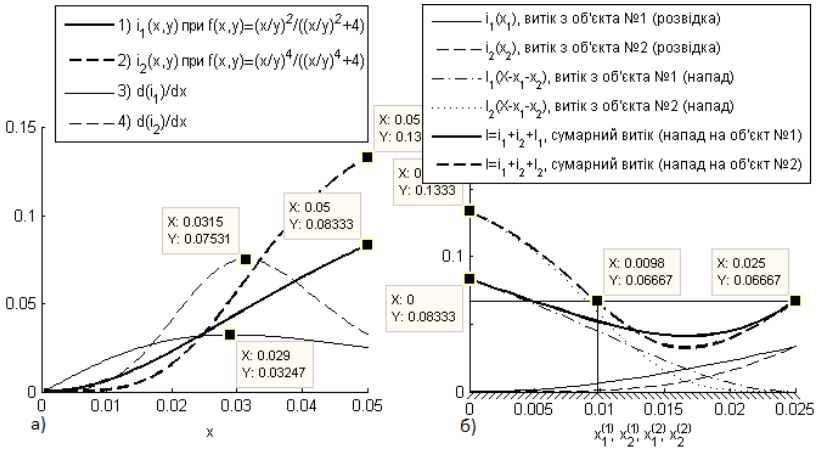



Рис.8.2

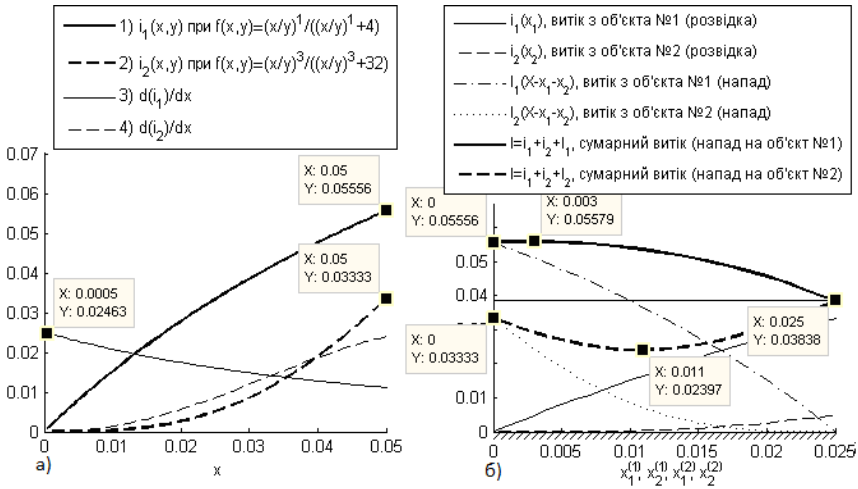



Рис.8.3

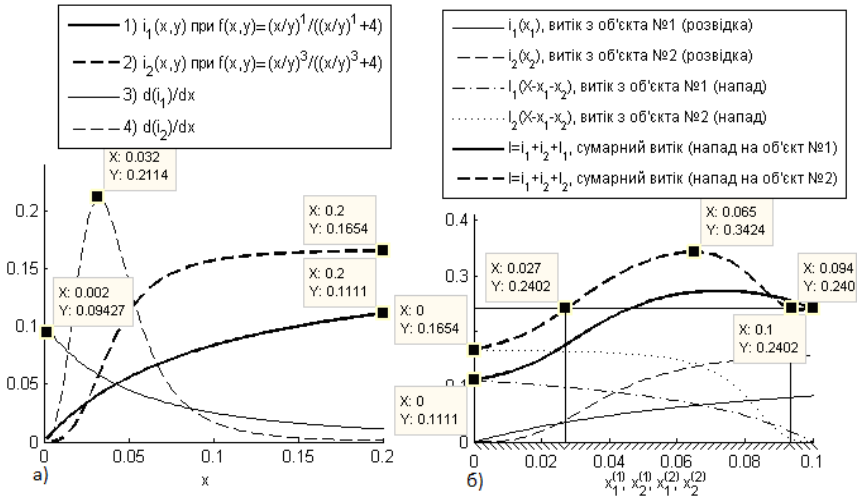



Рис.8.4

Одержані результати можна пояснити наступним чином.

Якщо розвідка проводиться в зоні низької динамічної вразливості (при $y \gtrsim 0$), то збільшення ресурсів розвідки не приносить бажаного результату, оскільки збільшення кількості

вилученої інформації під час розвідки не може компенсувати її зменшення під час витоку (крутизна кривої $f_2(\tilde{x})$ на рис.8.1,2 в зоні витоку при значних $x$ значно більша, ніж в зоні розвідки при $x \gtrsim 0$). В результаті сумарний витік зі зростанням $x$ зменшується – аж до точки, коли розвідка вийде на дільницю залежності $f_2(\tilde{x})$ з високою крутизною ($x = 0{,}0017$ на рис.8.1); фактично розвідка і витік при цьому міняються місцями.

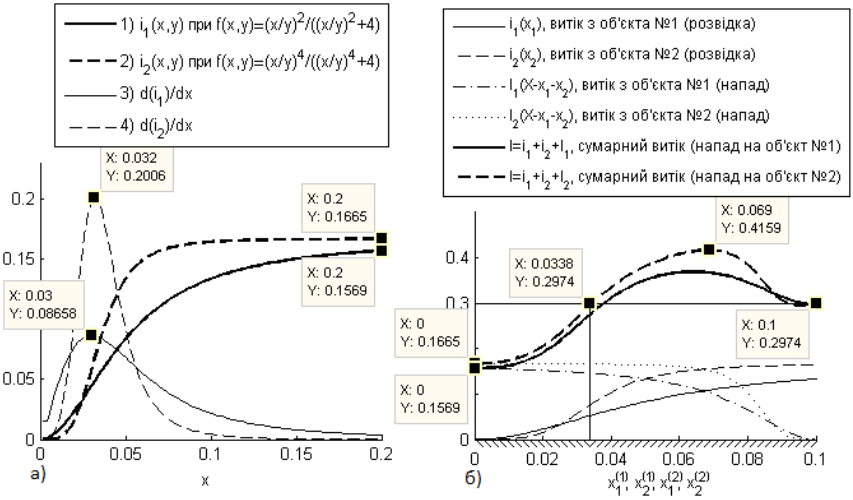



Рис.8.5

Цей висновок підтверджує і наступний, дещо несподіваний результат: при аномально низькій вразливості і малих ресурсах розвідки за критерієм порівняння кількості вилученої інформації розвідка стає доцільною при всіх $x$. На рис.8.3 зменшення вразливості досягнуто за рахунок збільшення значення $c$ в функції $f_2(\tilde{x})$ з 4 до 32. Доцільність розвідки при цьому обумовлена не збільшенням кількості інформації, вилученої в результаті проведення розвідки, а її зменшенням у відсутності розвідки (з $i = 0{,}067$ на рис.8.1,б до $i = 0{,}038$ на рис.8.3,б). Це пояснюється тим, що суцільна зростаюча крива $i(\tilde{x})$ на рис.8.1,б вигнулась і стала спадаючою на рис.8.3,б.

З сказаного випливає висновок, що остаточне рішення про доцільність проведення розвідки слід приймати, враховуючи не тільки співвідношення кількості інформації, вилученої без розвідки та з розвідкою, а й відношення кількості вилученої інформації в обох випадках до затрачених ресурсів, іншими словами – рентабельність.

Перевага варіанта з розвідкою (рис.8.3) над іншими зберігається у всьому інтервалі зміни $x^{(1)}$ – суцільна жирна лінія проходить вище горизонтальної лінії. Цей висновок справедливий у випадку, коли після розвідки буде прийнято правильне рішення, і всі ресурси направлені на перший об'єкт. При помилковому рішенні (всі ресурси направлені на другий об'єкт – штрихова жирна лінія) варіант з розвідкою виявляється гіршим за всі інші (тобто зосередження всіх ресурсів на одному з об'єктів або їх рівномірному розподілу між об'єктами). Це свідчить про важливість як проведення розвідки, так і правильного тлумачення її результатів.

При достатній загальній кількості ресурсів виділяти малу кількість ресурсів на розвідку немає сенсу, оскільки витік відбувається на положистій дільниці залежностей $f(\tilde{x})$ (рис.8.4,5), і зекономлені на розвідці ресурси не дають помітного збільшення $i(\tilde{x})$. З цієї причини зона доцільності розвідки на рис.8.4,5 переміщується на дільницю значних $x$.

Зазначимо, що, якщо два об'єкти описуються близькими зо формою функціями з яскраво вираженою нелінійністю (рис.8.6), то розвідка стає недоцільною – в зоні $x \gtrsim 0$ $i(x)$ близька до нуля і не може дати помітний внесок у сумарний витік. Цей висновок не є наслідком загального принципу, який твердить, що чим ближчі за характеристиками об'єкти, тим менше користі можна отримати від розвідки. В нашому випадку основну роль грає крутизна залежностей $f(\tilde{x})$ в робочих зонах. Це підтверджує рис.8.7, де схожі за властивостями об'єкти описуються залежностями $f(\tilde{x})$ з високою крутизною в зоні розвідки – і розвідка стає доцільною.

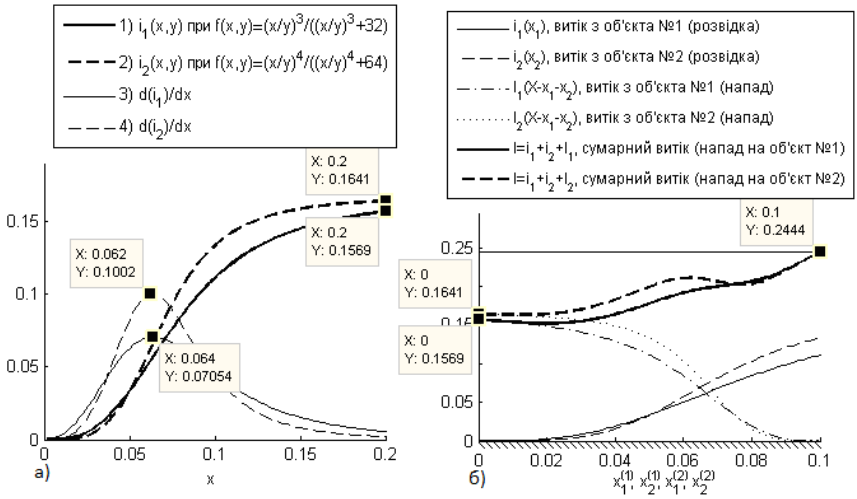



Рис.8.6

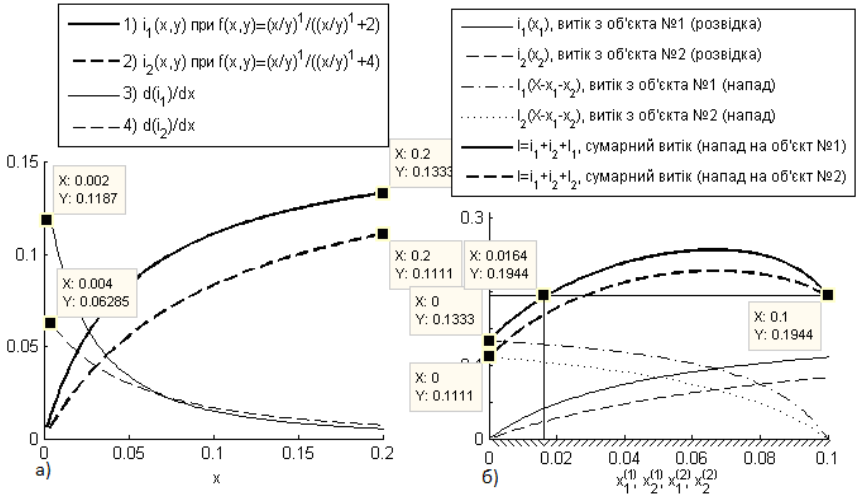



Рис.8.7

Оптимальні значення ресурсів, які необхідно виділяти на розвідку – $x_0^{(1)}$ і на витік інформації – $x_0^{(2)}$, досягаються в інтервалі значень $x$, де зростання функції $f(\tilde{x})$ уповільнюється (похідна $f'(\tilde{x})$ зменшується і згодом прямує до нуля). На рис.8.4

максимальне значення $i_{2\max}(\tilde{x}) = 0{,}3424$ для залежності $f_2(\tilde{x})$ (штрихова лінія) досягається при $x_0^{(1)} = 0{,}065$, що визначає величину $x_0^{(2)} = X - 2x_0^{(1)} = 0{,}2 - 0{,}124 = 0{,}076$. Ці обидва значення $x_0^{(1)}$ і $x_0^{(2)}$ знаходяться в інтервалі, де штрихова лінія на лівій частині рис.8.4 виходить на положисту дільницю – подальші інвестиції неефективні. Такий самий висновок можна зробити з рис.8.5. Розглядаючи приведений приклад в зворотному напрямку, можна визначити величину $X = 2x_1^{(1)} + x_0^{(2)}$, перевищення якої недоцільне. При використання функцій рис.8.4 $2x_0^{(1)} = 0{,}124$, і, як видно з положення штрихової лінії, $x^{(2)}$ бажано вибрати в інтервалі $0{,}07..0{,}1$, звідки $X = 0{,}194..0{,}224$. Задане в прямому розрахунку значення $X = 0{,}2$ лежить в цьому інтервалі, що свідчить про те, що зазначений рівень і розподіл ресурсів близькі до оптимального.

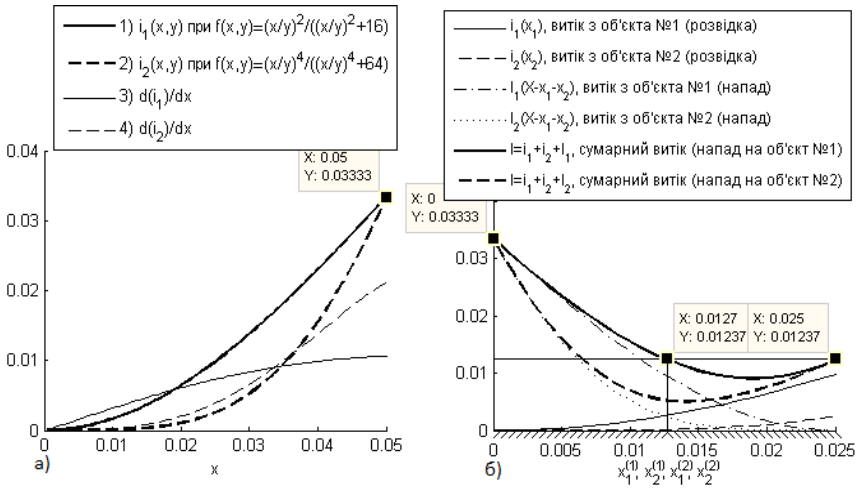



Рис.8.8

Приведені висновки в загальній своїй частині підтверджують розрахунки, виконані на основі широковідомої моделі Гордона-

Лоеба [3]. Ця модель ґрунтується на введеному понятті вразливості $\upsilon$, яка трактується як імовірність проникнення в об'єкт при відсутності інвестицій, спрямованих в його захист. Для визначення імовірності порушення безпеки при внесенні інвестицій $y$ введено два класи функцій:

$$S^I(y,\upsilon)=\frac{\upsilon}{(\alpha y+1)^\beta} \text{ та } S^{II}(y,\upsilon)=\upsilon^{\alpha y+1}.$$

Функція $S^I(y,\upsilon)$ принципово не відрізняється від використаних нами степеневих функцій, які можна записати як $f(x,y)=\frac{a}{1+c\left(y/x\right)^n}$, а при $a=\upsilon, c=\alpha, n=\beta=1, x=1$ ці функції повністю співпадають.

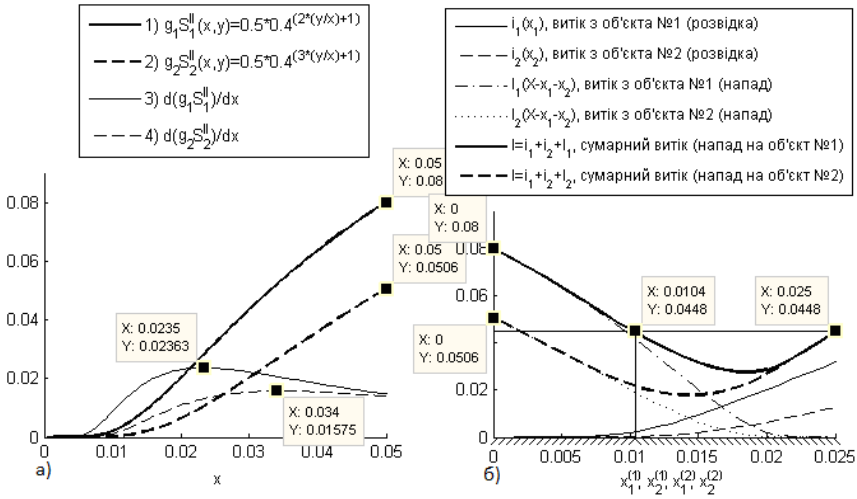



Рис.8.9

Розглядаючи імовірність $S(y,\upsilon)$ порушення безпеки як частку вилученої інформації, розрахуємо функцію $S^{II}(y/x,\upsilon)$ при низькому рівні ресурсів нападу ( $x=0..0{,}05$ ). Результати зображені

на рис.8.9. Подібну картину одержуємо при використанні степеневих функцій $f(x,y)$ (рис.8.2). На рис.8.10 використані функції різних класів – $S^I(y/x,\upsilon)$ та $S^{II}(y/x,\upsilon)$, для двох об'єктів при високому рівні ресурсів нападу ($x = 0..0{,}2$). Зображені залежності схожі на рис.8.4,5 і підтверджують приведені висновки.

Зазначимо, що одержання максимальної кількості вилученої інформації не є головною метою розвідки. Проте збільшення цієї величини дозволяє скласти більш повну картину системи захисту і тому покращує показники розвідки. При цьому слід мати на увазі, що низькі значення вилучення інформації під час розвідки не обов'язково спричиняють такі ж низькі величини під час вилучення (може бути навпаки).

Відзначимо ще декілька показників, які необхідно враховувати при прийнятті рішення про доцільність проведення розвідки. Перший з них – це інтервал значень $x$, в якому розвідка доцільна. Збільшення цього інтервалу зменшує ризик непродуктивних витрат, коли ми виходимо за його межі. Другий – це ступінь перевищення оптимального значення $i(\tilde{x})$, яке досягається при проведенні розвідки над значенням $i(\tilde{x})$ при її відсутності. Третій – ступінь близькості кривих, які зображають сумарний витік при нападі на об'єкт №1 і на об'єкт №2. Два останніх відношення характеризують рівень стійкості прийнятої методики до зміни її складових. І нарешті такий показник, як рентабельність загальних витрат, котрі можуть вирости внаслідок потреб на розвідку.

Відзначимо ще одну деталь, яку необхідно враховувати при проведенні розвідки: одна точка $f(\tilde{x})$, яку ми одержуємо в результаті розвідки не дає можливості передбачити форму всієї залежності. Для цього необхідно мати принаймні дві точки. Друга спроба може дати ще й побічний позитивний ефект: розподіл ресурсів на декілька спроб може пересунути робочу точку на дільницю залежності $f(\tilde{x})$ з великою крутизною, що збільшує сумарний результат. Ця можливість обумовлена тим, що зі зростанням $x$ $f(\tilde{x})$ зменшує свою крутизну, що є відбитком

відомого економічного закону про зменшення граничної норми прибутку.

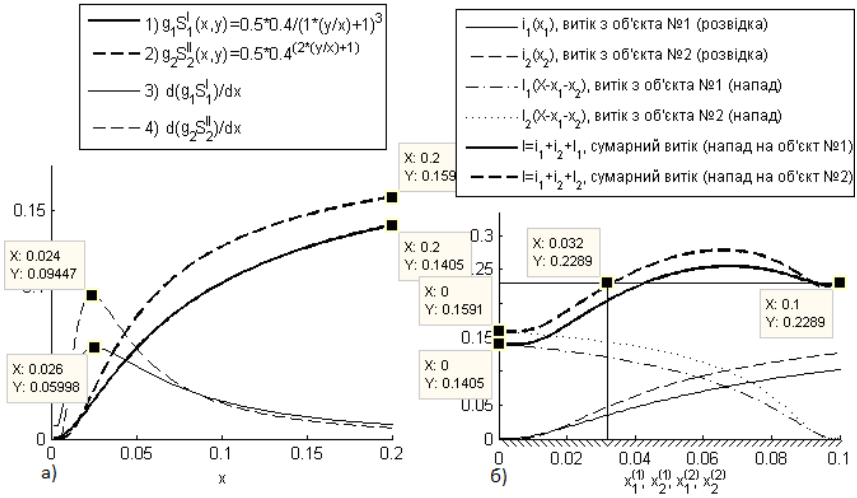



Рис.8.10

Як зазначено вище, оптимальні з точки зору одержання $i_{\max}$ значення ресурсів, які слід направляти на розвідку і на вилучення інформації, знаходяться на дільниці залежності $f(\tilde{x})$, де її зростання уповільнене, і $f'(\tilde{x})$ починає різко зменшуватись. Ця умова, з одного боку, підкреслює значення виду функції $f(\tilde{x})$, що характеризує динамічну вразливість об'єкта, з другого боку – ставить вимоги до необхідної кількості ресурсів нападу. При зміні одного з цих показників розвідка з недоцільної може переходити в доцільну і навпаки. Таким чином, питання про доцільність розвідки зводиться до одного з основних питань, які виникають при побудові математичної моделі – визначення складових і параметрів цільової функції, які в максимальній степені відповідають характеристикам реальних об'єктів. Це питання потребує свого дослідження та розв'язання.

### Контрольні питання

1. Роль розвідки в інформаційній сфері.
2. Чинники, які впливають на доцільність проведення розвідки.
3. Ефективність розвідки, її залежності від параметрів і характеристик математичної моделі.
4. Умови досягнення максимальної ефективності розвідки.
5. Особливості оптимального розподілу ресурсів при проведенні розвідки в системі з двох об'єктів.

## IX. ПРОДУКТИВНІСТЬ ІНВЕСТИЦІЙ В ІНФОРМАЦІЙНУ БЕЗПЕКУ

### 9.1 Поняття продуктивності

В моделі Гордона-Лоеба [22] введено поняття вразливості $\upsilon$, яка залежить від інвестицій $y$ в захист інформації (параметром в цю залежність входить початкова вразливість, котра визначається як $\upsilon$ при $y = 0$).

В [24] продуктивність інформаційної безпеки поділяється на два показники: продуктивність зменшення вразливості і продуктивність зменшення загрози. Перший з цих показників задається виразом $\upsilon^{\alpha y+1}$, а другий - $t^{\beta y+1}$, де $t$ - імовірність загрози, а $\alpha$ і $\beta$ - міри продуктивності обох типів. Рішення оптимізаційної задачі дозволяє знайти $y^0$ - розмір оптимальних інвестицій, при якому прибуток від інвестування досягає максимуму. Простір продуктивності за параметрами $\alpha$ і $\beta$ ділиться на зони, в одних з них $y^0 = 0$, в інших визначається виразами, які випливають з розв'язку задачі.

В [13] кількість вилученої інформації $i(x, y)$ при здійсненому нападі визначається через співвідношення ресурсів нападу і захисту - $x$ і, відповідно, $y$:

(9.1)

$$i(x,y) = q(x,y) \cdot f(x,y),$$

де $q(x,y)$ - імовірність виділення нападом ресурсів $x$ при заданому рівні ресурсів $y$; $f(x,y)$ - частка вилученої інформації при заданому співвідношенні $x$ і $y$;кількість інформації на об'єкті дорівнює$g=1$.

Використовуючи підхід [24], можемо вважати, що $q(x,y)$ пов'язано зі зменшенням загрози, а $f(x,y)$ - зі зменшенням вразливості при внесенні інвестицій $y$ в захист. В [13] запропоновані можливі варіанти цих залежностей. Враховуючи, що в (3) $x$ і $y$ входять у вигляді співвідношення $\frac{y}{x}$, введемо нову змінну $\tilde{y} = \frac{y}{x}$ і оберемо залежності $f(\tilde{y})$ і $q(\tilde{y})$ у такому вигляді (рис.9.1):

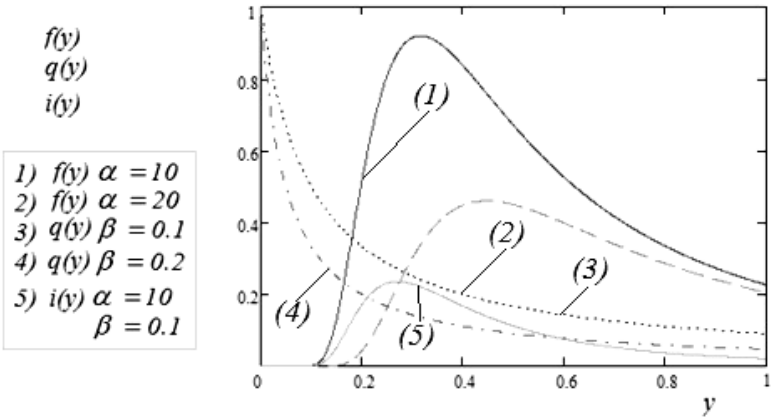



Рис. 9.1Форма залежностей $f(y)$, $q(y)$

$$f(\tilde{y}) = \frac{a}{1 + c\tilde{y}}, \tag{9.2}$$

$$q(\tilde{y}) = N \frac{1}{\tilde{y}^2} e^{-\frac{h^2}{\tilde{y}^2}}, \tag{9.3}$$

де $a$, $c$ і $h$ - параметри, а $N$ - коефіцієнт нормування.

Для спрощення запису надалі покладаємо $x=1$ та $\tilde{y}=y$. Використовуючи ці позначення, представимо (9.2) та (9.3) у вигляді:

$$f(y)=\frac{a}{1+\alpha y}, \tag{9.4}$$

$$q(y)=\frac{N}{y^2}e^{-\frac{\beta}{y^2}}. \tag{9.5}$$

З точки зору продуктивності інвестицій вважатимемо, що $\alpha=c$ є міра продуктивності зменшення вразливості, а $\beta=h^2$ - міра продуктивності зменшення загрози. Інтервали можливих значень $\alpha$ *і* $\beta$ встановлено з наступних міркувань. З (6) випливає*:* при*y=1* $f(y)=\frac{a}{1+\alpha}$. Вважатимемо, що при $\tilde{y}=\frac{y}{x}=1$ повинно бути $f<1$ (інакше систему захисту будемо вважати неефективною). Звідси випливає, що $\alpha>1$ (оскільки $a\le 1$). Можливі значення параметру $\beta$ встановимо, враховуючи, що $\beta=h^2$. Вважаємо, що ресурси нападу для успішної атаки при ефективній системі захисту повинні перевищувати ресурси захисту. Отже, $\beta$ лежить в межах: $0<\beta<1$.

## 9.2 Розрахунок продуктивності

Розглянемо варіанти систем, які відрізняються видами залежностей *f(y), q(y)*і визначимо границі зон продуктивності.

1.Побудуємо функцію, яка визначає прибуток від інвестицій (рис.9.2,а):

$$b(x,y)=i(x,0)-i(x,y)-y, \tag{9.6}$$

або, використовуючи (9.4), (9.5):

$$b(y)=1-\frac{N}{y^2}e^{-\frac{\beta}{y^2}}\cdot\frac{a}{1+\alpha y}-y, \tag{9.7}$$

де $i(x,0)=1$ - відносна кількість інформації, вилученої при $y=0$. З умови $\frac{db}{dy}=0$ одержуємо вираз для знаходження оптимального значення $y^0$:

$$b'(y)=\frac{N}{y^3}\cdot\frac{1}{1+\alpha y}\cdot(2-2\beta\cdot\frac{1}{y^2}+\frac{\alpha y}{1+\alpha y})\cdot e^{-\frac{\beta}{y^2}}-1=0. \qquad (9.8)$$

Дослідимо отриману функцію за допомогою програмних засобів математичного моделювання. На рис.9.2,б зображено хід похідної $b'(y)$ (10) при двох варіантах значень $\alpha_1=5$, $\beta_1=0.05$ (суцільна крива) і $\alpha_2=6$, $\beta_2=0.1$ (штрихова). Значення параметрів $\alpha$ і $\beta$ вибрані так, щоб проілюструвати різні ситуації у визначенні мір продуктивності.

Точка перетину суцільної кривої з віссю $y$ є точкою максимуму. Наприклад, точка К на кривій рис. 9.3,а відповідає точці перетину з віссю $y$ кривої 1 з параметрами $\alpha_1=5$, $\beta_1=0.05$ на рис. 9.2,б. При зміні мір продуктивності $\alpha$ та $\beta$ форма кривої $b'(y)$ (рис. 9.2,б) змінюється, і при значеннях параметрів $\alpha=6$ та $\beta=0.1$ крива стає дотичною до осі $y$ (пунктирна лінія). Ці значення $\alpha$ та $\beta$ (точка $L$ на рис. 9.3,а) є граничними – такими, що обмежують зону існування значень $y^0$. Граничні значення показників $\alpha$ та $\beta$ формують спадаючу криву на рис. 9.3,а.

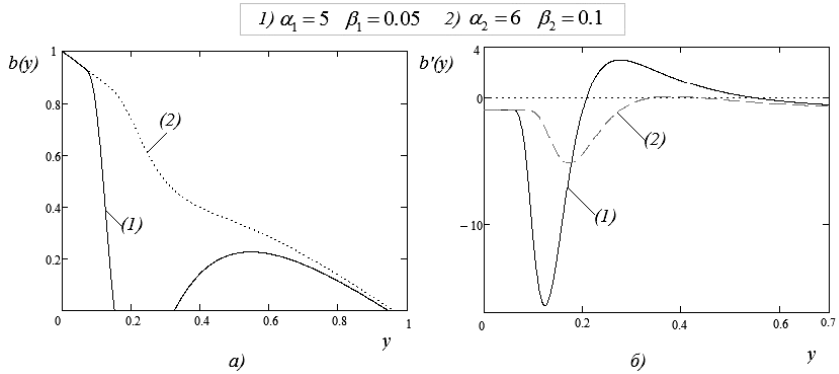



Рис. 9.2 Хід залежностей $b(y)$ та $b'(y)$

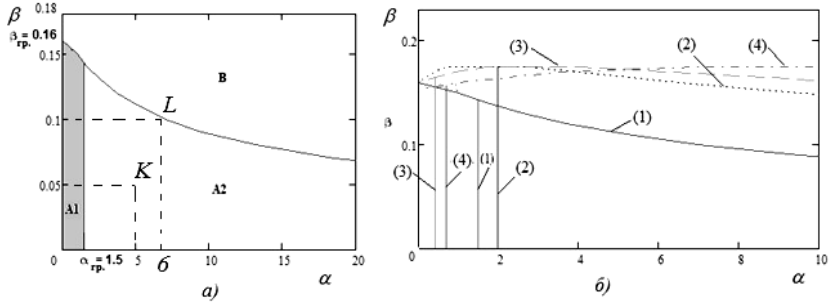



Рис. 9.3 Простір продуктивності при різних варіантах цільової функції

Простір продуктивності на рис. 9.3,а поділяється на три зони. А1 - малі $\alpha$ та малі $\beta$, А2 - великі $\alpha$ та малі $\beta$, В - великі $\beta$. Зауважимо, що $\alpha$ характеризує вразливість об'єкта (чим менше $\alpha$, тим більша вразливість $f$), а $\beta$ - відносну кількість ресурсів нападу (чим більше $\beta$, тим більше $\tilde{y}_m$, менше відношення $\frac{x}{y}$ і при сталому $x$ - більша кількість ресурсів захисту $y$). А1 - зона збитковості, в якій інвестування позбавлене сенсу, оскільки простіше відшкодувати збитки від втрати інформації, ніж вносити інвестиції в об'єкт з високою вразливістю. В зоні В при зростанні $\alpha$ (зменшення вразливості) граничне значення, при яких інвестиції

вже стають доцільними зменшується (гранична крива між зонами А2 і В є спадною).

2. Залежність $f(y)$(6) змінюємо з дрібно-лінійної на дрібно-квадратичну, залишаючи $q(y)$ незмінною. В результаті замість (9.7), (9.8) маємо:

$$b(y)=1-\frac{N}{y^2}\cdot e^{-\frac{\beta}{y^2}}\cdot\frac{1}{1+\alpha y^2}-y,$$
$$b'(y)=\frac{2N}{y^3}e^{-\frac{\beta}{y^2}}\frac{1}{1+\alpha y^2}(1-\frac{\beta}{y^2}+\frac{\alpha y^2}{1+\alpha y^2})-1. \qquad (9.9)$$

3. Залежність $q(y)$ змінюємо з розподілу Максвела на розподіл Релея, залишаючи $f(y)$ незмінною:

$$b(y)=1-\frac{N}{y}\cdot e^{-\frac{\beta}{y^2}}\cdot\frac{1}{1+\alpha y}-y,$$
$$b'(y)=\frac{N}{y^2}\cdot e^{-\frac{\beta}{y^2}}\cdot\frac{1}{1+\alpha y}(1-\frac{2\beta}{y^2}+\frac{ay^2}{1+ay^2})-1. \qquad (9.10)$$

4. Використовуємо обидві змінені залежності $f(y)$ і $q(y)$ і в результаті маємо:

$$b(y)=1-\frac{N}{y}\cdot e^{-\frac{\beta}{y}}\cdot\frac{1}{1+\alpha y^2}-y,$$
$$b'(y)=\frac{N}{y^2}\cdot e^{-\frac{\beta}{y^2}}\cdot\frac{1}{1+\alpha y^2}(1-\frac{2\beta}{y^2}+\frac{2ay^2}{1+ay^2})-1. \qquad (9.11)$$

На рис.9.3,б зображений простір продуктивності для різних варіантів функції $b(y)$. Кожна границя складається з двох частин – вертикальної прямої, яка визначається граничним значенням $a_{гр.}$ і кривою $\beta(\alpha)$ при $\alpha>\alpha_{гр.}$. При заміні $f(y)=\frac{1}{1+\alpha y}$ на

$f(y)=\dfrac{1}{1+\alpha y^2}$ (перехід від варіанту 1 до варіанту 2), дещо збільшується зона нульових інвестицій А1 (гранична пряма між зонами А1 та А2 трохи зміщується вправо) і розширюється зона А2 за рахунок підйому границі між зонами А2 та В. При заміні $q(y)=N\dfrac{1}{y^2}e^{-\frac{\beta}{y^2}}$ на $q(y)=\dfrac{N}{y}\cdot e^{-\frac{\beta}{y^2}}$ (варіант 3), відбувається значне зменшення зони нульових інвестицій та відносно невелике збільшення зони А2. Таким чином, для варіанту 2 ми маємо найбільшу зону нульових інвестицій, для варіанту 3 – найменшу, для варіанту 4 – найбільшу зону А2, для варіанту 1 - найменшу.

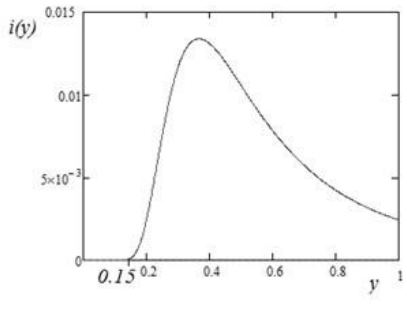



Рис.9.4. Залежність вилученої інформації від інвестованих ресурсів $y$

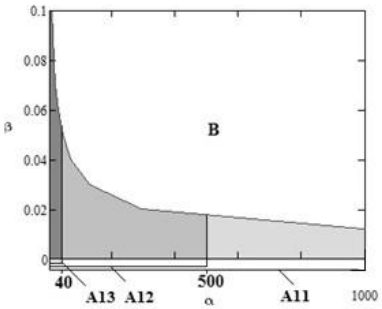



Рис.9.5. Зони продуктивності для системи з трьох об'єктів

Розглянемо як приклад інформаційну систему, яка містить три об'єкти з об'ємами інформації g1 = 0.2, g2 = 0.3, g3 = 0.5. Міри продуктивності для об'єктів обрані такими, щоб оптимальні значення інвестицій знаходились в різних зонах. Для першого об'єкта $\alpha_1=30,\ \beta_1=0.02$ - це зона А1. Для другого - $\alpha_2=100,\ \beta_2=0.2$ (зона В). Для третього - $\alpha_3=100\ \beta_3=0.01$ (зона А2). Границі зон нульових інвестицій визначаються з умови $b(y)=0$ (9) і становлять: $\alpha_{гр.1}=2000,\ \alpha_{гр.2}=500,\ \alpha_{гр.3}=40.$ З (10) знаходимо оптимальний розмір інвестицій: для першого об'єкту

оптимум знаходиться в зоні А1 і становить $y_1^0 = 0,$ для третього об'єкту оптимум – в зоні А2 і становить $y_3^0 = 0.239$. Для другого об'єкту оптимум знаходиться в зоні В і визначається з (3), за допомогою рис.9.4, і лежить в інтервалі $y_2^0 \in (0;0.15)$. На рис.9.5 побудовано зони продуктивності для представленого варіанту.

На рисунку позначено: А11 – зона нульових інвестицій для першого об'єкту, А12 – для другого, А13 – для третього.

Приведені розрахунки дозволяють розробити стратегію вибору оптимального значення ресурсів захисту в залежності від заданих мір продуктивності зниження загрози і зниження вразливості.

### Контрольні питання

1. Види продуктивності інвестицій і розрахунок їх показників.
2. Зв'язок показників продуктивності з вразливістю об'єкта.
3. Зони продуктивності, їх сутність і визначення границь.
4. Варіанти завдання цільової функції, які визначають границі зон продуктивності.
5. Зв'язок показників продуктивності між собою.

## X. ДИНАМІЧНЕ УПРАВЛІННЯ РЕСУРСАМИ ЗАХИСТУ ІНФОРМАЦІЇ

### 10.1 Метод Белмана

Динамічне управління ресурсами направлене на підвищення ефективності їх використання в умовах динамічної взаємодії зловмисника і захисника. При моделюванні цього процесу основним математичним інструментом побудови алгоритму динамічного управління ресурсами є динамічне програмування – метод оптимізації, який має застосування до операцій, в котрих процес прийняття рішень може бути розбитий на етапи (кроки). Такі операції називають багатокроковими. В інформаційній сфері подібні ситуації виникають, зокрема, при пошуку оптимального управління ресурсами в просторі і в часі, тобто між об'єктами захисту інформації на протязі низки послідовних періодів.

Обчислювальна схема динамічного програмування ґрунтується на використанні принципу оптимальності Р.Белмана[3]: яким би не був стан системи після певного числа кроків, на найближчому кроці слід вибирати управління так, щоб разом з оптимальним управлінням на всіх наступних кроках воно приводило до сумарного оптимального виграшу.

Введемо позначення:

$s_k$ – стан системи після $k$-го кроку;

$x_k$– управління на $k$-му кроці;

$i_k(s_{k-1}, x_k)$ – цільова функція (виграш) $k$-го кроку.

В нашому прикладі$s_k$ – це відстані, які ми можемо подолати після $k$-го кроку:

$x_k$– варіанти $k$-го кроку;

$i_k$ – відстані, які ми можемо пройти на $k$-му кроці.

Позначимо через $i_k^*(s_{k-1})$максимум цільової функції, який досягається за умови, що перед $k$-м кроком система була в $(k-1)$-му стані і управління на $k$-му кроці було оптимальним:

$$i_n^*(s_{n-1}) = \max_{\{i_n\}} i_n(s_{n-1}, x_n)$$

Ця величина називається умовним максимумом цільової функції на $n$-му кроці.

Відповідно до принципу оптимальності $x_k$вибираються так, щоб досягався максимум суми приведених величин на всіх $n$-кроках:

$$i_n^*(s_{n-1}) = max_{\{x_k\}}\{i_k(s_{k-1}, x_k) + i_k^*\}, \quad k = n-1, n-2, .., 2, 1.$$

Типовою задачею динамічного програмування є логістична задача, в якій необхідно вибрати найбільш економний (в найпростішому випадку - найкоротший) маршрут між двома кінцевими пунктами,з'єднані між собою великою кількістю проміжних пунктів. Задачу можна розв'язати методом перебору – порівнюючи різні багатокрокові маршрути (крок – це подолання відстані між сусідніми пунктами). Вибір сусіднього пункту є кроком управління. Таким чином, ми приходимо до задачі

[3]Р. Белман (1920 – 1984) – американський математик.

оптимізації багатокрокового управління. Проблема полягає в тому, що оптимізація окремого кроку (тобто вибір найближчого сусіднього пункту) зовсім не гарантує одержання результуючого оптимального маршруту. Крокове управління повинно обиратись з врахуванням можливих наступних кроків, а вибір найближчого пункту може відвести нас від оптимального маршруту. Необхідність передбачення всіх можливих продовжень маршруту при великій кількості проміжних пунктів може привести до значних обчислювальних ускладнень. Однак в описаній процедурі продовження маршруту є один виняток, позбавлений труднощів вибору. Це останній крок, в якому нема потреби передбачати наступні кроки (їх просто не існує). Це приводить до необхідності формування процесу динамічного програмування в зворотному напрямку – від останнього кроку до першого.

Для ілюстрації принципу оптимальності розглянемо близьку до нашого напрямку класичну задачу про розподіл ресурсів між підприємствами. Нашою метою є пошук оптимального розподілу, який забезпечує одержання сумарного максимального прибутку. Прибуток $i_k$ ($k$ - номер підприємства) залежить від внесених ресурсів$x$, причому ці залежності для різних підприємств мають різний характер і ілюструються табл.10.1, в якій приведені значення$i_k(x)$ для чотирьох підприємств у випадку, коли загальна сума ресурсів становить 5 одиниць і розміри внесків в підприємство кратні одній одиниці.

Таблиця10.1.

Прибутки $i_k$ підприємств в залежності від внесених ресурсів $x$

| $x$ \ $i_k$ | $i_1$ | $i_2$ | $i_3$ | $i_4$ |
|---|---|---|---|---|
| 1 | 8 | 6 | 5 | 6 |
| 2 | 10 | 9 | 6 | 7 |
| 3 | 11 | 11 | 9 | 8 |
| 4 | 13 | 12 | 11 | 13 |
| 5 | 18 | 15 | 17 | 16 |

Рішення починаємо з кінця. Перший варіант розподілу – всі ресурси вкладаємо в одне підприємство. Очевидно, це буде перше підприємство і прибуток від вкладання всіх коштів в це

підприємство становить: $i^{(1)}(x)=i_1(5)=18$ (верхній індекс – номер варіанту, нижній – номер підприємства).

Другий варіант: 4 одиниці ресурсів вкладаємо в одне підприємство і 1 одиницю – в інше. Об'єкти, які забезпечують найбільший прибуток при 4 внесених одинцях – перший і четвертий, при 1 одиниці – перший (інші варіанти не розглядаємо). Отже, кращий варіант:

$$i^{(2)}(x)=i_4(4)+ i_1(1)=13+8=21.$$

Третій варіант: 3 одиниці ресурсів направляємо на одне з підприємств, 2 одиниці – в інше. З аналогічних міркувань обираємо друге і перше підприємства:

$$i^{(3)}(x)=i_2(3)+ i_1(2)=11+10=21.$$

Четвертий варіант: 3 одиниці направляємо на одне з підприємств і по одній – на два інших:

$$i^{(4)}(x)=i_2(3)+ i_1(1)+ i_4(1)=11+8+6=25.$$

П'ятий варіант: 2 одиниці направляємо на одне з підприємств і по одній – на три інших. Одержуємо:

$$i^{(5)}(x)=i_2(2)+ i_1(1)+ i_3(1)+ i_4(1)=9+8+5+6=28.$$

Отже, оптимальним варіантом є п'ятий, який забезпечує найбільший прибуток.

Приведений приклад дозволяє виявити дві основні відмінності принципу оптимальності:

1) багатокрокова оптимізаційна задача замінюється на низку однокрокових;

2) при порівнянні проміжних маршрутів відкидаються заздалегідь невигідні, зменшуючи загальну кількість варіантів.

Спробуємо «спроектувати» розглянуту задачу на інформаційну сферу. Поставимо завдання: знайти оптимальний розподіл ресурсів *x*нападу на три об'єкти, при якому досягається максимальне значення $i(x)$ вилученої інформації. Об'єкти мають однакову кількість інформації $g_k = \frac{1}{3}$, однакову кількість ресурсів захисту $y=1$ і відрізняються лише вразливістю. В цих умовах:

$$i_k(x) = g_k p_k f_k(x) = \frac{1}{3} \cdot 1 \cdot f_k(x) = \frac{1}{3} f_k(x)$$

Для того, щоб визначити функції $i_k(x)$ для трьох об'єктів, задамо залежності$f_k(x)$ у вигляді:

$$f_1(x)=\frac{x}{x+2};\quad f_2(x)=\frac{x^3}{x^3+10};\quad f_3(x)=\frac{x}{4}.$$

В цих виразах враховано, що *y=1* для всіх об'єктів, тобто під *x* розуміємо відношення *x/y* на кожному об'єкті. Провівши розрахунки в інтервалі значень *x=0,5…2,5*, заповнюємо таблицю 10.2.

Таблиця 10.2.

Частка вилученої інформації в залежності від затрачених ресурсів нападу

| N | $f_k(x)$ / $x$ | $f_1(x)$ | $f_2(x)$ | $f_3(x)$ |
|---|---|---|---|---|
| 1 | 0,5 | 0,20 | 0,01 | 0,12 |
| 2 | 1,0 | 0,33 | 0,09 | 0,25 |
| 3 | 1,5 | 0,43 | 0,25 | 0,38 |
| 4 | 2,0 | 0,50 | 0,44 | 0,50 |
| 5 | 2,5 | 0,56 | 0,61 | 0,63 |

Розглядаючи варіанти поділу загальної суми *x=2,5* на різну кількість об'єктів, визначимо для кожного поділу оптимальні об'єкти і в результаті одержимо:

$$i^{(1)}(x)=\frac{1}{3}f_3(2.5)=\frac{1}{3}\cdot 0.63=0.21$$

(всі ресурси – в третій об'єкт);

$$i^{(2)}(x)=\frac{1}{3}[f_2(2.0)+f_1(0.5)]=\frac{1}{3}(0.44+0.20)=0.21$$

(поділ ресурсів на частини – 4+1);

$$i^{(3)}(x)=\frac{1}{3}[f_3(1.5)+f_1(1.0)]=\frac{1}{3}(0.38+0.33)=0.24$$

(поділ 3+2);

$$i^{(4)}(x)=\frac{1}{3}[f_2(1.5)+f_1(0.5)+f_3(0.5)]=\frac{1}{3}(0.25+0.20+$$

$$+0.12)=0.19$$

(поділ 3+1+1);

$$i^{(5)}(x)=\frac{1}{3}[\,f_1(1.0)+f_3(1.0)+f_2(0.5)]=\frac{1}{3}(0.33+0.25+$$

$$+0.01)=0.20$$

(поділ 2+2+1).

Оптимальним є розподіл $i^{(3)}(x)$ з найбільшим значенням $i(x)=0.24$.Відповідно до цього варіанту на 1-ий об'єкт необхідно направити 1 одиницю ресурсів нападу, на 3-ій – 1,5 одиниць, а на 2-ий нічого не направляти.

Розглянемо тепер задачу оптимізації розподілу ресурсів захисту. Покладаючи $x=1$ для кожного об'єкта, перейдемо від приведених залежностей $f_k(x)$ до залежностей$f_k(y)$:

$$f_1(y)=\frac{1}{1+2y};\quad f_2(y)=\frac{1}{1+10y^3};\quad f_3(y)=\frac{1}{4y}.$$

Результати розрахунків в інтервалі $y=0.25\ldots1.25$ (що відповідає оберненим значенням $x=4\ldots0.8$, близьким, на наш погляд, до реальності) приведені в табл.10.3.

Таблиця 10.3.

Частка втраченої інформації в залежності від вкладених ресурсів захисту

| N | $f_k(y)$ / $y$ | $f_1(y)$ | $f_2(y)$ | $f_3(y)$ |
|---|---|---|---|---|
| 1 | 0,25 | 0,80 | 0,94 | 1,00 |
| 2 | 0,50 | 0,67 | 0,88 | 0,67 |
| 3 | 0,75 | 0,57 | 0,70 | 0,44 |
| 4 | 1,00 | 0,50 | 0,33 | 0,33 |
| 5 | 1,25 | 0,44 | 0,20 | 0,27 |

Приведемо результати розрахунків загальної кількості втраченої інформації для різних варіантів поділу сумарних ресурсів на частини при оптимальному виборі об'єктів інвестування (оптимум тепер визначається мінімальним значенням $i(y)$). При розрахунках вважатимемо, що з незахищених об'єктів при нападі вилучається вся інформація: $i_k(0)=1$. Оскільки дії нападу нам невідомі, вважатимемо, що імовірності нападу на всі об'єкти

однакові, $p_k = \frac{1}{3}$. Інформація, як і раніше, розподілена між об'єктами порівну: $g_k = \frac{g}{3} = \frac{1}{3}$.

Тоді $i(y) = \sum_{k=1}^{3} g_k p_k f_k(y) = \frac{1}{9} \sum_{k=1}^{3} f_k(y)$. В результаті маємо:

$$i^{(1)}(y) = \frac{1}{9}[i_1(0) + i_2(1.25) + i_3(0)] = \frac{1}{9}(1 + 0.05 + 1) = 0.23$$

$$i^{(2)}(y) = \frac{1}{9}[i_1(0.25) + i_2(1.00) + i_3(0)] = \frac{1}{9}(0.67 + 0.09 + +1) = 0.20$$

$$i^{(3)}(y) = \frac{1}{9}[i_1(0.50) + i_2(0.75) + i_3(0)] = \frac{1}{9}(0.50 + 0.19 + +1) = 0.19$$

$$i^{(4)}(y) = \frac{1}{9}[i_1(0.25) + i_2(0.25) + i_3(0.75)] = \frac{1}{9}(0.67 + 0.83 + +0.33) = 0.20$$

$$i^{(5)}(y) = \frac{1}{9}[i_1(0.50) + i_2(0.25) + i_3(0.50)] = \frac{1}{9}(0.50 + 0.43 + +1) = 0.21$$

Оптимальним є розподіл з найменшим значенням $i(y)$, а саме $i^{(3)}(y)=0.19$. Цей варіант передбачає виділення на 1-ий об'єкт 0,50 одиниць ресурсів, на 2-ий – 0,75 одиниць, а 3-ій залишити без захисту.

Зазначимо різницю в постановці задачі двох останніх прикладів, яка приводить до різних числових коефіцієнтів у виразах для $i(x)$: $\frac{1}{3}$ при розгляді дій нападу і $\frac{1}{9}$ при розподілі ресурсів захисту. У першому випадку вразливості об'єктів, відображені в табл.10.2, відомі нападу, тому він має можливість вести націлені атаки, і задача зводиться до порівняння обмеженої

кількості варіантів розподілу. В другому випадку (табл.10.3) захист не має відомостей про наміри суперника, задача стає стохастичною, і необхідно враховувати імовірності нападів на різні об'єкти.

## 10.2 Аналіз підходів до динамічного управління ресурсами захисту інформації

Із зростанням рівня загроз безпеці організації і постійного обмеження бюджету для компанії дуже важливо визначити, в які аспекти інформаційної безпеки потрібно вкласти інвестиції і в якій кількості. Враховуючи той факт, що забезпечити безпеку організації на 100% неможливо, за ціль ставиться досягнути ефективного і розумного інвестування з виконанням поставленої задачі – досягнути максимально можливого рівня безпеки при заданій кількості ресурсів чи забезпечити необхідний рівень безпеки при мінімальних затратах. Тому для осіб, що приймають рішення (менеджер чи експертна група) дуже важливо вирахувати оптимальний обсяг інвестицій.

Важливим завданням менеджменту інформаційної безпеки є управління ресурсами, направлене на досягнення максимальної ефективності їх використання. Показником ефективності може бути кількість вилученої інформації, прибуток за рахунок зменшення витоку інформації або рентабельність, яка визначається як відношення прибутку до розміру інвестицій. Оптимізації підлягають наступні величини:

1) розмір інвестицій Y;
2) розподіл інвестицій по об’єктах $\{y_k\}$, $\sum_k y_k = Y$;
3) момент $t^0$інвестування.

Існування оптимуму відносно моменту інвестування можна пояснити наступними міркуваннями. Затримка в інвестуванні в умовах послідовних атак, звичайно, приведе до певних втрат. Проте попередній розподіл ресурсів, коли ще не проявилась націленість суперника, може виявитись неефективним і привести до ще більших втрат. Тому настає момент $t^0$, який визначається пороговим значенням $i^0$кількості вилученої інформації, при

настанні якого стає доцільним виділення певної кількості ресурсів захисту $y^0$ і певного розподілу $\{ y_k^0 \}$ їх між об'єктами.

При розробці методики управління ресурсами можна виділити два підходи: проактивний – коли ресурси захисту вносяться ще до першого нападу зловмисника, тобто здійснюється попереднє інвестування, і реактивний або адаптивний, коли вся або більша частина коштів вноситься захистом після здійснення першої атаки зловмисником. В залежності від наявності інформації про можливі загрози (або степінь невизначеності відносно дій нападника) в проактивному виді захисту можуть використовуватися різні варіанти розподілу: при достатній обізнаності захисту про можливі цілі нападу ресурси розподіляються динамічно, тобто з розрахунком пропорцій внесення інвестицій на кожний з об'єктів, знаходженням декількох варіантів розподілу і вибором оптимального; за браком інформації захисник буде вимушений розподілити свої кошти рівномірно між всіма об'єктами для забезпечення мінімального захисту. Вибір варіанту розподілу при достатній обізнаності захисту очевидний – динамічний розподіл відповідно до наявних загроз. Але виникає питання, чи доцільно у другому випадку, коли є певна невизначеність у діях зловмисника, застосовувати рівномірний розподіл. Адаптивний підхід до управління має на меті вирішення саме цього питання.

### 10.3 Адаптивний підхід як вид динамічного управління

Необхідність адаптивного управління ресурсами інформаційної безпеки обумовлена такими причинами:

1) невизначеністю відносно дій суперника, а саме направленістю його зусиль по вилученню інформації і масштабом цих зусиль;

2) зміною з часом як внутрішніх, так і зовнішніх умов протистояння – стану інформаційної системи (вартості інформації і її розподілу між об'єктами), направленості атак суперника, появою нових суперників тощо.

3) обмеженістю ресурсів, виділених на інформаційну безпеку.

Розглянемо доцільність застосування адаптивного методу інвестування в інформаційну безпеку.

Використаємо введену раніше цільову функцію, яка визначає кількість вилученої інформації:

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

де :*i(x,y)* – загальна кількість вилученої інформації;

*x,y* – ресурси нападу і захисту відповідно;

$k=\overline{1,l}$ – номер об'єкта;

$g_k$ – кількість інформації на *k*-му об'єкті;

$p_k(x,y)$ – імовірність нападу на *k*-ий об'єкт;

$q_k(x,y)$ – імовірність виділення ресурсів *x* при нападі на *k*-ий об'єкт;

$f_k(x,y)$ – частка вилученої інформації з *k*-го об'єкта.

Залежності $f_k(x,y)$ можуть приймати такий вигляд :

$$f_k(x,y)=\frac{\left(x/y\right)^{n_k}}{a_k\left(x/y\right)^{n_k}+c_k} \qquad (2)$$

Параметри $a_k$, $c_k$, $n_k$визначають форму залежності для кожного об'єкта.

Як перший приклад оберемо залежності$f(x,y)$однаковими у такій формі:

$$f(x,y)=\frac{x/y}{2\left(x/y\right)+8} \qquad (3)$$

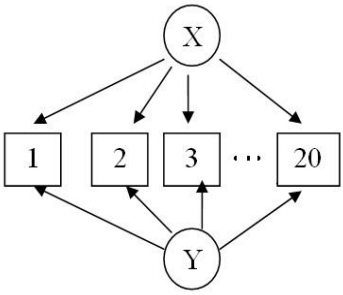
Сталі величин (3) вибрані такими, що при рівних ресурсах нападу і захисту ($x/y=1$) кількість вилученої інформації з одного об'єкта становить 10% ($f(x,y)=0.10$) – це значення вважаємо близьким до реальності, а при відсутності захисту ($x/y\to\infty$) вилучається максимум 50% ( $f(x,y)\to 0.5$ ), що визначається

початковою вразливістю об’єкта. Вважаємо, що після нападів інформація поповнюється.

Параметри для проведення дослідження вибираємо наступні: кількість об’єктів *l=20*,кількість інформації на них однакова:$g_k=g$, напад здійснюється на 2 об’єкти, при чому проводяться 3 атаки. Загальна кількість ресурсів *Y=1, X=2*. Розглядаємо спрощений варіант, коли $p_k(x,y)=1$ і $q_k(x,y)=1$. В силу однаковості об’єктів отримуємо:

$$i(x,y)=\sum_{k=1}^{2} g_k f_k(x,y)=2g\cdot f(x,y) \qquad (4)$$

Порівняємо два види розподілу - проактивний і реактивний (адаптивний). Оскільки за умовою націленість нападу невідома, в проактивному захисті ресурси розподіляємо порівну між всіма об’єктами (рис. 9.1). В адаптивному інвестиції вносяться на 2 об'єкти після першого нападу (рис. 9.2).



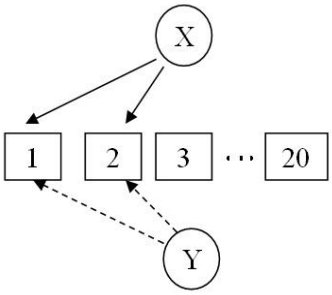




Рис.10.1 Рис.10.2

Проактивний:

$$y_k=y=\frac{Y}{l}=\frac{1}{20}=0.05;\quad x_k=x=\frac{X}{2}=1$$

Втрачена інформація після кожної атаки:

$$i_1=i_2=i_3=2g\cdot\frac{\frac{1}{0.05}}{\frac{2}{0.05}+8}=0.833g$$

Після трьох атак втрати інформації становлять:$i_{акт}=2.49g$

Адаптивний:

1) перша атака*x=1, y=0:* $i_1 = 2g \cdot 0.5 = g$
2) друга атака*x=1, y=1/2:* $i_2 == 2g \cdot 0.167 = 0.333g$
3) третя атака*x=1, y=1/2:* $i_3 = i_2 = 2g \cdot 0.167 = 0.333g$ .

Після трьох нападів втрати інформації становлять:

$$i_{адап} = (1 + 0.333 + 0.333)g = 1.67g$$

Перевага адаптивного підходу виражається відношенням:

$$E = \frac{I_{акт}}{I_{адап}} = \frac{2.49}{1.67} = 1.5.$$

### 10.4 Вплив умов нападу ефективність адаптивного підходу

Опрацьовуються декілька варіантів складу системи та нападу, які відрізняються загальною кількістю об'єктів і кількістю атакованих об'єктів, а також кількістю атак. Проілюструємо доцільність застосування адаптивного підходу за допомогою показника ефективності $E = \frac{I_{àêò}}{I_{àäàï}}$ для різних функцій$f_k(x,y)$, які характеризують вразливості об'єктів (табл.10.4-10.10). Графіки функцій $f_k(x,y)$ приведені на рис.10.3.

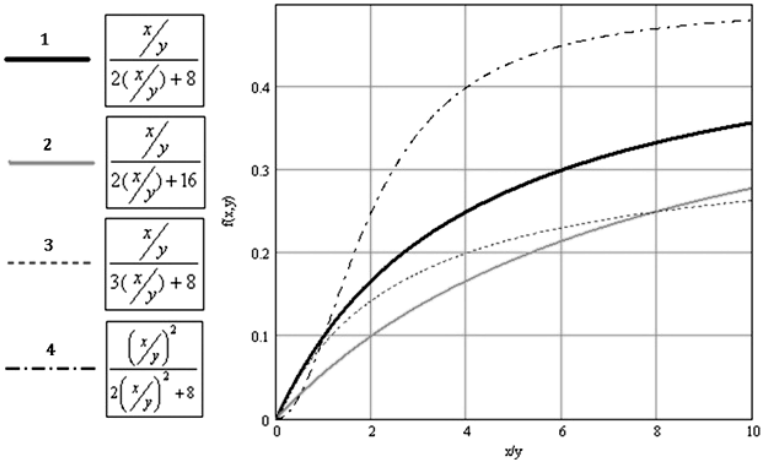



Рис. 10.3.

Таблиця10.4.

20 однакових об’єктів, 3 атаки, напад на *k* об’єктів

| $E = \frac{Iакт}{Iадап}$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| № | $f_k(x,)$ \ $k$ | 2 | 3 | 5 | 6 | 8 | 9 | 10 | 14 | 15 |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1,5 | 1,385 | 1,2 | 1,125 | 1 | 0,947 | 0,9 | 0,75 | 0,72 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 1,531 | 1,339 | 1,071 | 0,974 | 0,824 | 0,765 | 0,714 | 0,564 | 0,536 |
| 3. | $\frac{x/y}{3(x/y)+8}$ | 1,432 | 1,346 | 1,212 | 1,154 | 1,054 | 1,01 | 0,969 | 0,836 | 0,808 |
| 4. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,485 | 1,467 | 1,412 | 1,376 | 1,293 | 1,247 | 1,2 | 1,007 | 0,96 |

Затінені клітинки відносяться до ситуацій, коли адаптивний підхід недоцільний (*E<1*).

З таблиць 10.4 та 10.5 можна зробити наступні висновки:

Адаптивний підхід доцільніше використовувати при великій кількості атак (2 та більше). Це ілюструється більшою граничною кількістю об’єктів при 3 нападах, ніж при 2. Для кожної з функцій маємо: функція 1 – 8 та 5 відповідно, функція 2 – 5 та 3, функція 3 – 9 та 6, функція 4 – 14 та 11. Таким чином, розрахунки показують, що раціональніше застосовувати даний вид розподілу при здійсненні послідовних нападів на систему:

для $\frac{\frac{x}{y}}{2\frac{x}{y}+8}$: 8 та 5; для $\frac{\frac{x}{y}}{2\frac{x}{y}+16}$: 5 та 3;

для $\frac{\frac{x}{y}}{3\frac{x}{y}+8}$: 9 та 6; для $\frac{\left(\frac{x}{y}\right)^2}{2\left(\frac{x}{y}\right)^2+8}$: 14 та 11.

Таблиця 10.5.

20 об’єктів, 2 атаки, напад на *k* об’єктів

| $E = \frac{Iакт}{Iадап}$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| № | $f_k(x,y)$ \ $k$ | 2 | 3 | 4 | 5 | 6 | 10 | 11 | 12 | 15 |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1,25 | 1,154 | 1,071 | 1 | 0,938 | 0,75 | 0.714 | 0.682 | 0,6 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 1,19 | 1,042 | 0,926 | 0,833 | 0,758 | 0,556 | 0.521 | 0.49 | 0,417 |
| 3. | $\frac{x/y}{3(x/y)+8}$ | 1,235 | 1,167 | 1,105 | 1,05 | 1 | 0,84 | 0.808 | 0.778 | 0,7 |
| 4. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,32 | 1,304 | 1,282 | 1,255 | 1,223 | 1,067 | 1,024 | 0,98 | 0,853 |

Ефективність підходу при різній загальній кількості об’єктів проілюстровано в таблицях 10.4,10.6-10.8. З приведених результатів можна зробити наступні висновки: при різній загальній кількості об’єктів ефективність адаптивного підходу залишається сталою, що демонструє гранична кількість об’єктів, при якій ще доцільно використовувати даний метод. Для найбільш стійких систем (функція 4) адаптивний підхід залишається ефективним при нападі на 60-70% об’єктів, для більш вразливих систем (функції 1 та 3) - 40%, і для систем з найбільшою вразливістю (функція 2) – 20-25%.

Таблиця 10.6.

15 об’єктів, 3 атаки, напад на *k* об’єктів

| $E = \frac{Iакт}{Iадап}$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| № | $f_k(x,y)$ \ $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 10 | 11 |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1,421 | 1,286 | 1,174 | 1,08 | 1 | 0,931 | 0,771 | 0,73 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 1,398 | 1,19 | 1,037 | 0,918 | 0,824 | 0,748 | 0,584 | 0,545 |
| 3. | $\frac{x/y}{3(x/y)+8}$ | 1,372 | 1,275 | 1,192 | 1,118 | 1,054 | 0,996 | 0,855 | 0,817 |
| 4. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,474 | 1,442 | 1,4 | 1,35 | 1,293 | 1,232 | 1,038 | 0,975 |

Таблиця 10.7.

10 об’єктів, 3 атаки, напад на *k* об’єктів

| $E = \frac{Iакт}{Iадап}$ | | | | | | |
|---|---|---|---|---|---|---|
| № | $k$ / $f_k(x,y)$ | 2 | 3 | 4 | 5 | 7 |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1,286 | 1,125 | 1 | 0,9 | 0,75 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 1,19 | 0,974 | 0,824 | 0,714 | 0.564 |
| 3. | $\frac{x/y}{3(x/y)+8}$ | 1,275 | 1.154 | 1,054 | 0,969 | 0.836 |
| 4. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,442 | 1,376 | 1,293 | 1,2 | 1,007 |

Розглянуті варіанти одержані при умові наявності на всіх об’єктах однакової кількості інформації. Розглянемо тепер випадок, коли система має різні об’єкти. Система складається з 10 об’єктів. Інформація розподіляється:

$g_1=0.3$, $g_2=0.2$, $g_k = 0.0625$ $(k = \overline{3,10})$, вся інформація $g=1$.

Таблиця 10.8.

5 об’єктів, 3 атаки, напад на *k* об’єктів

| $E = \frac{Iакт}{Iадап}$ | | | | |
|---|---|---|---|---|
| № | $k$ / $f_k(x,y)$ | 2 | 3 | 4 |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1 | 0,818 | 0.692 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 0,824 | 0,63 | 0.51 |
| 3. | $\frac{x/y}{3(x/y)+8}$ | 1,05 | 0,897 | 0.782 |
| 4. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,293 | 1,103 | 0,915 |

Напад здійснюється на перші 2 об’єкти. Умови проактивного розподілу незмінні. Адаптивний розподіл має декілька варіантів, в яких об’єм резервування (відкладення коштів на інвестування після першої атаки) становить:

1) 50%; 2) 70%; 3) 80%; 4) 100%.

Ця частина коштів вноситься після першої атаки і розподіляється між двома об’єктами пропорційно кількості інформації. Результати для різної кількості атак показані в табл. 10.9 та 10.10 (клітини, затемнені діагональною лінією, показують оптимальний варіант динамічного підходу).

В попередніх розрахунках ми розглядали ситуації, в яких взагалі потрібно розподіляти ресурси безпеки у відповідності до адаптивного підходу. Таблиці 10.9,10 ілюструють вже різні варіанти цього підходу. Перед менеджером ставиться задача в залежності від характеристик системи, що потребує захисту, вибрати раціональний спосіб розподілу ресурсів. Звичайно, для кожної системи розподіл буде різний, але, відповідаючи на питання ефективності резервування в адаптивному підході, ми отримуємо однакові результати майже для всіх приведених функцій, а відповідно і для систем з різною вразливістю. Дані таблиця доводить найбільшу ефективність застосування адаптивного підходу з резервуванням – стримуванням інвестицій до атаки (внесенням 0% інвестицій) і направленням їх на об’єкти в повному обсязі вже після першого нападу. Але, все ж виявляється і доцільність реалізації інших варіантів внесення інвестицій – із внесенням певної частини коштів вже перед першою атакою (це підтверджує той факт, що показники ефективності мають значення більше одиниці).

Вибір оптимального варіанту резервування залежить від вибраної функції вразливості в кожному конкретному випадку. Вразливість може оцінюватись за допомогою різноманітного набору характеристик і параметрів. В даному випадку використовується кількість вилученої з об’єкта інформації. Функції кількості вилученої інформації зображено на рис.10.4.

Найкращий варіант резервування при двох та трьох атаках для одних систем однаковий – варіант повного резервування, для систем, що описуються першими трьома функціями, найбільш ефективний варіант внесення певного відсотка інвестицій до атаки.

На діаграмах 10.5,6 ефективність застосування певних варіантів резервування представлена графічно відповідно до ступеня вразливості об’єктів.

Таблиця 10.9.

10 різних об’єктів, 3 атаки, напад на 2 об’єкти

| $E = \frac{Iакт}{Iадап}$ | | | | | |
|---|---|---|---|---|---|
| № | $k$ / $f_k(x,y)$ | 50% | 70% | 80% | 100% |
| 1. | $\frac{x/y}{2(x/y)+8}$ | 1,175 | 1,247 | 1,27 | **1,29** |
| 2. | $\frac{x/y}{2(x/y)+16}$ | 1,212 | **1,254** | 1,249 | 1,193 |
| 3. | $\frac{x/y}{2(x/y)+24}$ | 1,216 | **1,22** | 1,187 | 1,062 |
| 4. | $\frac{x/y}{3(x/y)+8}$ | 1,144 | 1,215 | 1,242 | **1,281** |
| 5. | $\frac{x/y}{4(x/y)+8}$ | 1,121 | 1,186 | 1,213 | **1,256** |
| 6. | $\frac{x/y}{10(x/y)+8}$ | 1,061 | 1,099 | 1,116 | **1,149** |
| 7. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1,126 | 1,249 | 1,317 | **1,457** |
| 8. | $\frac{(x/y)^3}{2(x/y)^3+8}$ | 1,046 | 1,123 | 1,177 | **1,31** |

Характеристика вразливості об’єктів впливає на вибір варіантів резервування таким чином, що для систем з середньої вразливістю найбільш раціональніше буде внести 20-50% інвестицій до першого нападу, мінімально захистивши їх. Для систем, що зі значно меншою вразливістю такі інвестиції непотрібні, оскільки існує природна захищеність об’єктів, і менеджер виграє більше від повного резервування. Ті ж дії будуть раціональнішими і для систем з більшою від середнього показника вразливістю, оскільки внесення інвестицій в малій кількості не забезпечить повного захисту, ефективніше для менеджера буде прийняти певні витрати після першої атаки, а потім максимально захистити вразливі об'єкти. Таким чином, знаючи вразливість об'єктів, можна описати системи за допомогою однієї з функцій і визначити оптимальний варіант адаптивного підходу.

Таблиця 10.10.

10 різних об’єктів, 2 атаки, напад на 2 об’єкти

| $E = \frac{Iакт}{Iадап}$ | | | | | |
|---|---|---|---|---|---|
| № | $f_k(x,y)$ \ $k$ | 50% | 70% | 80% | 100% |
| 1. | $\frac{x/y}{2(x/y)+8}$ | **1.076** | 1.094 | 1.093 | 1.074 |
| 2. | $\frac{x/y}{2(x/y)+16}$ | **1.064** | 1.041 | 1.012 | 0.927 |
| 3. | $\frac{x/y}{2(x/y)+24}$ | **1,041** | 0,981 | 0,93 | 0,796 |
| 4. | $\frac{x/y}{3(x/y)+8}$ | 1.069 | 1.098 | 1.105 | **1.109** |
| 5. | $\frac{x/y}{4(x/y)+8}$ | 1,062 | 1,092 | 1,103 | **1,115** |
| 6. | $\frac{x/y}{10(x/y)+8}$ | 1,034 | 1,057 | 1,067 | **1,084** |
| 7. | $\frac{(x/y)^2}{2(x/y)^2+8}$ | 1.083 | 1.164 | 1.206 | **1,291** |
| 8. | $\frac{(x/y)^3}{2(x/y)^3+8}$ | 1,033 | 1,088 | 1,126 | **1,214** |

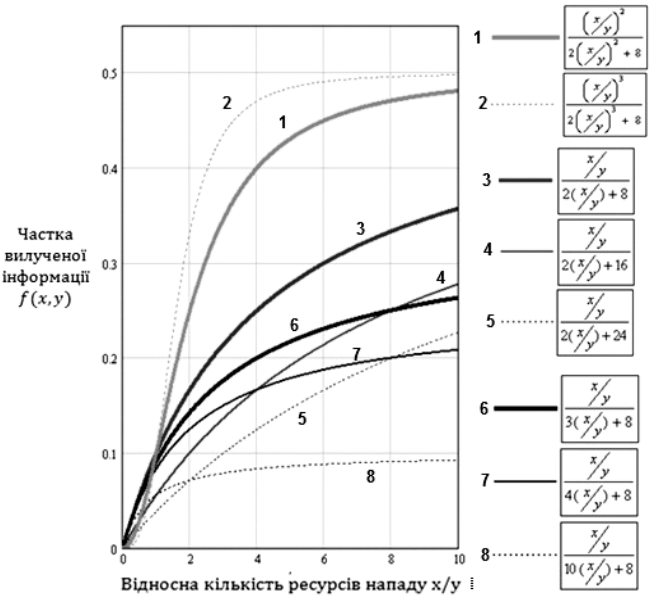



Рис. 10.4

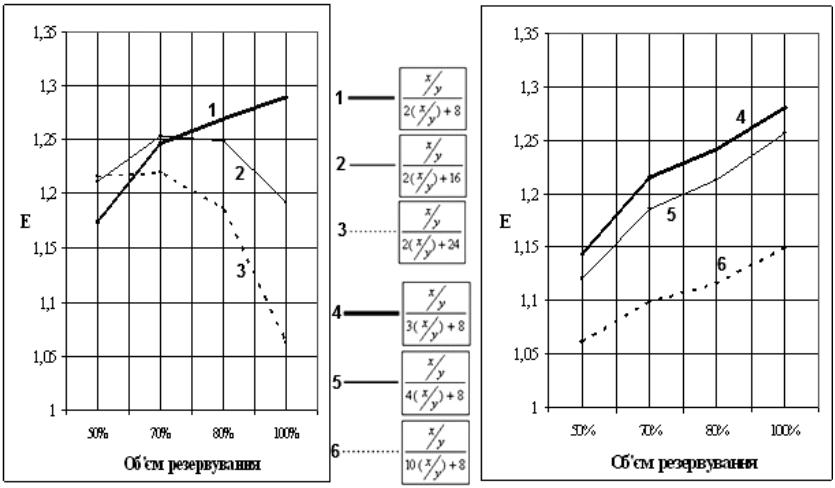



Рис10.5.

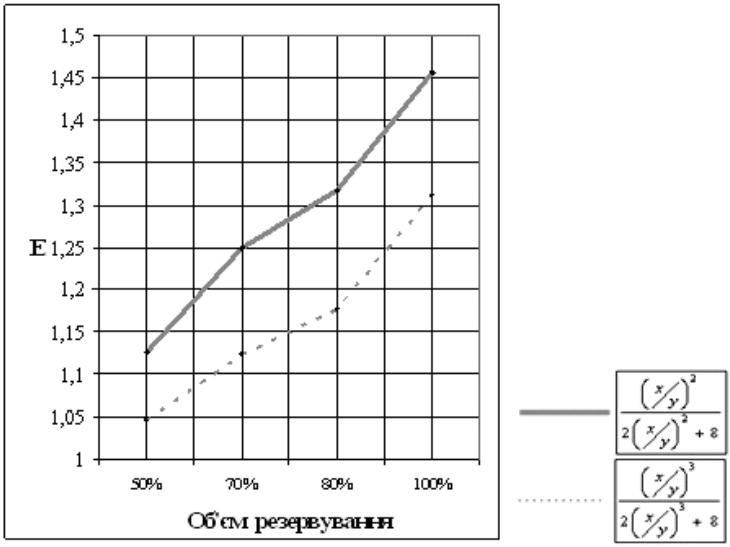



Рис.10.6

## Контрольні питання

1. Сутність принципу оптимальності Белмана і його застосування.
2. Порівняння підходів до динамічного управління ресурсами захисту інформації.
3. Методика розрахунку втрат інформації при проактивному і адаптивному розподілі ресурсів.
4. Доцільність адаптивного підходу і умови, за яких досягається його найбільша ефективність.

## XI. ОПТИМІЗАЦІЯ СУМАРНИХ ВТРАТ У СФЕРІ ЗАХИСТУ ІНФОРМАЦІЇ

### 11.1 Постановка задачі

Основним завданням економічного менеджменту інформаційної безпеки являється мінімізація можливих втрат інформації при одночасній мінімізації витрат на її захист. Ці два частинні критерії оптимальності суперечливі і не можуть бути виконані одночасно. Тому доводиться розглядати компромісний варіант, в якому зазначені показники враховуються з певними коефіцієнтами.

Цільову функцію, яка враховує втрати $I(x, y)$від витоку інформації і витрати$Y$на її захист, побудуємо у вигляді:

$$S(x,y) = \lambda I(x,y) + (1-\lambda)Y, \qquad (11.1)$$

де $\lambda$ і $1 - \lambda$ - вагові коефіцієнти відповідних величин.

Подібний підхід в застосуванні до інформаційної безпеки використано в [8], де цільова функція являє собою суму витрат на захист інформації і втрат від її сподіваного витоку. Метою дослідження в [8] є визначення розміру втрат, при якому цільова функція досягає мінімуму. Дослідженню підлягали окремі форми цільової функції,які в загальному вигляді відрізнялись швидкістю зміни сподіваних втрат від розміру витрат (в наших позначеннях це $I'(y)<0$).

Необхідно ввести вагові коефіцієнти для величин $I(x,y)$і$Y$,оскільки витрати на захист інформації і втрати від її витоку нерівнозначні для підприємства. Будемо оцінювати значення вагових коефіцієнтів $\lambda$ і $1 - \lambda$ по розміру небезпеки, яку справляють величини $I(x, y)$і$Y$на економічній стан підприємства, хоча зрозуміло, що встановити дійсний розмір втрат від витоку інформації в деяких випадках досить проблематично. Розрахувати вартість втрат можна лише в окремих випадках, наприклад, коли маємо справу з витратами на відновлення пошкодженої бази даних. Значно важче це зробити в умовах конкурентної боротьби, якщо в результаті вилучення інформації суперник одержує змогу захопити

додаткову частину ринку і використовувати цю можливість невизначений час, оскільки передбачити розвиток подій практично неможливо. В умовах стохастичної задачі, коли параметри розрахунку і функціональні залежності, які входять в розрахунок, визначаються лише з певною імовірністю, введення зазначених коефіцієнтів є доцільним.

## 11.2 Результати розрахунків

Переходячи до розрахунків, покладемо $p=1$ і пронормуємо всі величини в (11.1) до $g$:

$$s(x,y) = \lambda i(x,y) + (1-\lambda)y, \qquad (11.2)$$

де $s(x,y) = \frac{S(x,y)}{g}$, $i(x,y) = \frac{I(x,y)}{g}$, $y = \frac{Y}{g}$

На відміну від (11.1) в (11.2)всі величини відносні.

Нашим наступним кроком є визначення залежностей $q(x)$ і $f(x,y)$.В першому наближенні покладемо $q(x) = const = q$. Величину $q$ визначимо з умови $qx_m = 1$,де $x_m = 3$, звідки $q=1/3$. Тоді

$$s(x,y) = \frac{1}{3}\lambda f(x,y) + (1-\lambda)y. \qquad (11.3)$$

Введемо позначення:

$j = 1 - i$ – кількість захищеної інформації;

$b = j - y = 1 - i - y$– прибуток від інвестицій в захист інформації;

$r = b/y$ – рентабельність інвестицій;

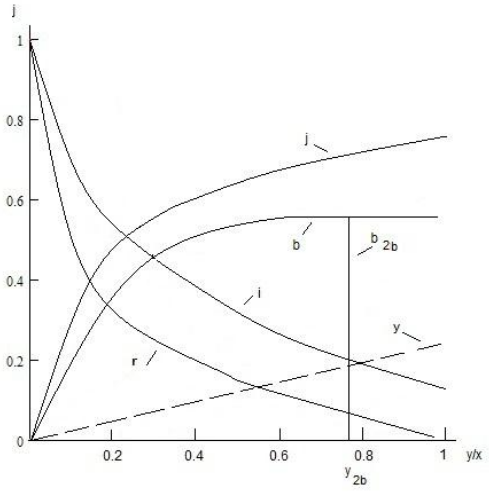
$y_{kb}$ – розмір інвестицій, при якому досягається максимальний прибуток;

$y_{kr}$ – розмір інвестицій, при якому досягається максимальна рентабельність;

$y_{ko}$ – оптимальне значення $y$, яке визначається з розрахованих значень$y_{kb}$, $y_{kr}$в залежності від критерію оптимальності.

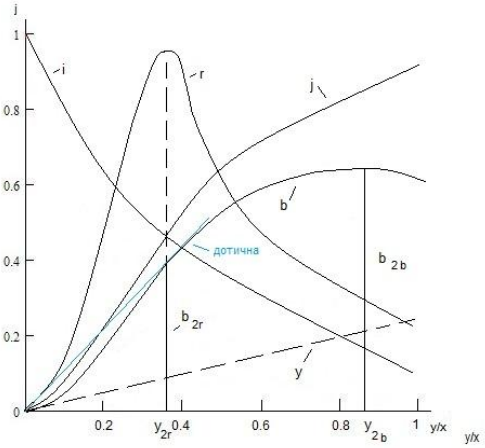
На рис. 11.1, 11.2 зображені залежності $i$, $j$, $b$,$r$від $y$для двох об'єктів, які відрізняються видом залежності $f(x,y)$ – для першого об'єкта вона має дрібно-лінійний характер, для другого – дрібно-нелінійний. Ця відмінність справляє суттєвий вплив на хід

залежностей$r(y)$ і положення точок$y_{1r}$, $y_{2r}$. Оскільки $r(y)$розраховуються як значення$tg\alpha$, де $\alpha-$ кут між відрізком прямої, проведеної до відповідної точки кривої $b(y)$ і віссю абсцис, то крива $r_1(y)$ буде спадною і досягне максимального значення при$y_{1r}=0$, а $r_2(y)$досягає максимуму в точці$y_{2r}<y_{2b}$. Ця точка визначається дотичною, проведеною з початку координат до кривої$b(y)$. Зазначені особливості обумовлені напрямком опуклості кривої$b(y)$ в області$y \cong 0$: для першого об'єкта вона направлена вгору, а для другого – вниз. Оптимальний розмір інвестицій $y_{k0}$для кожного об'єкта знаходиться в інтервалі між$y_{kr}$ і$y_{kb}$ та визначається в залежності від приорітетів, які надаються прибутковості і рентабельності. При дрібно-нелінійному характері залежності$f(x,y)$ величина $y_{k0}$може співпадати з $y_{kr}$, забезпечуючи максимальну рентабельність. При дрібно-лінійному характері функції $f(x,y)$максимальна рентабельність досягається при$y=0$ і $y_{k0}>y_{kr}$. Використовуючи приведену методику, можна при заданих значеннях$\lambda$ розрахувати комплексну цільову функцію$s(x,y)$, яка включає зазначені показники для кожного об'єкта, а потім і для всієї системи.




$$i_2(x,y) = g_2 f_2(x,y) = 0.5\frac{(x/y)^2}{(x/y)^2+8}$$

Рис.11.1 Залежності основних показників від розміру інвестицій при дробно-лінійному характері функції f(x,y)




$$i_1(x,y) = g_1 f_1(x,y) = 0.5 \frac{x/y}{x/y + 4}$$

Рис. 11.2 Залежності основних показників від розміру інвестицій при дробно -нелінійному характері функцій f(x,y)

## Контрольні питання

1. Вид багатоцільових функцій, їх складові.
2. Залежності основних економічних показників від розміру інвестицій при різних формах функції *f(x,y).*
3. Положення максимумів прибутку і рентабельності при різних формах складових цільової функції.

## XII. ЗАСТОСУВАННЯ НЕЧІТКОЇ ЛОГІКИ І ТЕОРІЇ НЕЧІТКИХ МНОЖИН В ЗАДАЧАХ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Оптимальний розподіл ресурсів між об'єктами захисту інформації є однією з важливих задач економічного менеджменту інформаційної безпеки. Оскільки пошук оптимального рішення ведеться в умовах невизначеності, здається доцільним розглянути можливість пошуку рішення на базі теорії нечітких множин, основи якої заклав Л. Заде[4] в 1965р. [26].

### 12.1 Основи теорії нечітких множин

Нечітка множина – це клас об'єктів, в якому немає різкої межі між тими об'єктами, які входять у цей клас, і тими, які в нього не входять. Більш точне визначення може бути сформульовано наступним чином.

Нехай $X = \{x\}-$ сукупність об'єктів (точок), що позначаються через $x$. Тоді нечітка множина $\tilde{A}$ в $X$ є сукупність упорядкованих пар

$$\tilde{A} = \{x, \mu_{\tilde{A}}(x)\}, \quad x \in X,$$

де $\mu_{\tilde{A}}(x)$ представляє собою ступінь належності $x$ до $\tilde{A}$,

а $\mu_{\tilde{A}} : X \to M -$ функція, що відображає $X$ в простір $M$, який називається простором належності.

Коли $M$ містить тільки дві точки *0* і *1*, $\tilde{A}$ є чіткою (точною) і його функція належності співпадає з характеристичною функцією чіткої множини.

У подальшому ми будемо припускати, що $M$ є інтервал [0,1], причому 0 і 1 представляють відповідно нижчий та вищий ступень належності. Таким чином, наше основне припущення полягає в тому, що нечітка множина $\tilde{A}$, незважаючи на нечіткість його

[4] Л. А. Заде (нар. в 1921 р.) – американський математик.

кордонів, може бути точно визначена шляхом зіставлення кожному об'єкту $X$ числа, що лежить між *0* і *1*, яке представляє ступінь його належності до $\tilde{A}$.

Для раціонального запису зручно було б мати засіб для вказування того, що нечітка множина $\tilde{A}$ отримана з чіткої множини $A$ за рахунок «розмиття» кордону множини $A$. Для цієї мети ми будемо використовувати хвилясту риску над символом (або символами), що визначають $\tilde{A}$. Символ ~ буде називатися оператором розмиття (*fuzzifier*). Перетин $\tilde{A}$ і $\tilde{B}$ позначається $\tilde{A} \cap \tilde{B}$ і визначається як найбільша нечітка множина, що міститься як в $\tilde{A}$, так і в $\tilde{B}$. Функція належності для $\tilde{A} \cap \tilde{B}$ визначається наступною рівністю:

$$\mu_{\tilde{A}\cap\tilde{B}}(x) = min\left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right), \quad x \in X,$$

де $min(a,b) = a$, якщо $a \le b$, і $min(a,b) = b$, якщо $a > b$.

Якщо використовувати замість символу $min$ знак кон'юнкції $\wedge$, можна переписати умову у більш простому вигляді:

$$\mu_{\tilde{A}\cap\tilde{B}} = \mu_{\tilde{A}} \wedge \mu_{\tilde{B}}.$$

Поняття об'єднання множин подібне поняттю перетину. Об'єднання $\tilde{A}$ і $\tilde{B}$ позначається $\tilde{A} \cup \tilde{B}$ і визначається як найменша нечітка множина, що містить як $\tilde{A}$, так і $\tilde{B}$. Функція належності для $\tilde{A} \cup \tilde{B}$ визначається співвідношенням

$$\mu_{\tilde{A}\cup\tilde{B}}(x) = max\left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right), \quad x \in X,$$

де $max(a,b) = a$, якщо $a \ge b$, і $max(a,b) = b$, якщо $a < b$.

Використовуючи замість символу $max$ знак диз'юнкції $\vee$, можна записати умову (5) у більш простому вигляді:

$$\mu_{\tilde{A}\cup\tilde{B}} = \mu_{\tilde{A}} \vee \mu_{\tilde{B}}.$$

У загальноприйнятому підході головними елементами процесу прийняття рішення є: а) множина альтернатив, б) множина обмежень, які необхідно враховувати при виборі між різними альтернативами та в) функція переваги, що ставить у відповідність

кожній альтернативі виграш (або програш), який буде отриманий в результаті вибору цієї альтернативи.

При розгляді цього процесу з більш загальних позицій прийняття рішень в нечітких умовах природною видається інша логічна схема, найважливішою рисою якої є симетрія по відношенню до цілей та обмежень [20]. Ця симетрія усуває розходження між цілями та обмеженнями і дозволяє досить просто сформувати на їх основі рішення.

Дійсно, нехай $X = \{x\}-$ задана множина альтернатив. Тоді нечітка ціль, або просто ціль, $\tilde{G}$ буде ототожнюватися з фіксованою нечіткою множиною $\tilde{G}$ в $X$. При звичайному підході функція переваги, що використовується в процесі прийняття рішення, служить для встановлення лінійної впорядкованості на множині альтернатив [16]. Очевидно, що функція належності $\mu_G(x)$ нечіткої цілі виконує ту ж задачу і, звичайно, може бути отримана з функції переваги за допомогою нормалізації, що зберігає встановлену лінійну впорядкованість. По суті, така нормалізація приводить до загального знаменника різні цілі й обмеження і дозволяє, таким чином, звертатися з ними однаковим чином. Як ми побачимо, це є важливим аргументом на користь того, щоб в якості одного з основних компонентів у логічній схемі прийняття рішень в нечітких умовах користуватися поняттям цілі, а не функції переваги.

Подібним же чином нечітке обмеження, або просто обмеження, $\tilde{C}$ у просторі $X$ визначається як деяка нечітка множина $\tilde{C}$ в $X$.

Нехай у просторі альтернатив $X$ задані нечітка ціль $\tilde{G}$ і нечітке обмеження $\tilde{C}$. Тоді нечітка множина $\tilde{D}$, утворена перетином $\tilde{G}$ і $\tilde{C}$, називається рішенням. У символічній формі

$$\tilde{D} = \tilde{G} \cap \tilde{C},$$

і відповідно

$$\mu_{\tilde{D}} = \mu_{\tilde{G}} \wedge \mu_{\tilde{C}}.$$

У більш загальному випадку, якщо є $n$ цілей і $m$ обмежень, то результуюче рішення визначається перетином всіх заданих цілей та обмежень, тобто

$$\tilde{D} = \tilde{G}_1 \cap \tilde{G}_2 \cap ... \cap \tilde{G}_n \cap \tilde{C}_1 \cap \tilde{C}_2 \cap ... \cap \tilde{C}_m,$$

і відповідно

$$\mu_{\tilde{D}} = \mu_{\tilde{G}_1} \wedge \mu_{\tilde{G}_2} \wedge ... \wedge \mu_{\tilde{G}_n} \wedge \mu_{\tilde{C}_1} \wedge \mu_{\tilde{C}_2} \wedge ... \wedge \mu_{\tilde{C}_m}.$$

Зауважимо, що в наведеному визначенні нечіткого рішення цілі й обмеження входять у вираз для $\tilde{D}$ абсолютно однаковим чином, що й доводить твердження про тотожність цілей та обмежень у сформульованій нами логічній схемі процесів прийняття рішень в нечітких умовах. У загальному випадку приймемо, що $\tilde{D}-$ нечітке рішення з функцією належності $\mu_{\tilde{D}}$.

У визначенні розпливчастого рішення $\tilde{D}$ як перетину або, в більш загальному сенсі, як злиття цілей і обмежень мається на увазі, що всі цілі й обмеження, що входять в $\tilde{D}$, мають однакову важливість. Однак трапляються ситуації, в яких деякі цілі і, можливо, деякі обмеження є більш важливими, ніж інші. У таких випадках рішення $\tilde{D}$ може бути виражене опуклою комбінацією цілей та обмежень з ваговими коефіцієнтами, що характеризують відносну важливість складових елементів. Таким чином, $\mu_{\tilde{D}}(x)$ може бути записано у вигляді:

$$\mu_{\tilde{D}}(x) = \sum_{i=1}^{n} \alpha_i(x) \cdot \mu_{\tilde{G}_i}(x) + \sum_{j=1}^{m} \beta_j(x) \cdot \mu_{\tilde{C}_j}(x), \qquad (12.1)$$

де $\alpha_i$ и $\beta_j -$ функції належності, такі, що $\sum_{i=1}^{n} \alpha_i(x) + \sum_{j=1}^{m} \beta_j(x) \equiv 1.$

З урахуванням цього обмеження функції $\alpha_i$ і $\beta_j$ можуть бути підібрані так, щоб передати відносну важливість цілей $\tilde{G}_1, \tilde{G}_2, ..., \tilde{G}_n$ та обмежень $\tilde{C}_1, \tilde{C}_2, ..., \tilde{C}_m$. Зокрема, якщо $m = n = 1,$

неважко перевірити, що з виразу *(1)* можна отримати будь-яку нечітку множину, що міститься в $\tilde{G} \cup \tilde{C}$ і включає $\tilde{G} \cap \tilde{C}$.

Досі ми обмежувалися розглядом ситуацій, в яких цілі і обмеження є нечіткими множинами в просторі альтернатив $X$. Практичний інтерес представляє більш загальний випадок, коли цілі й обмеження – нечіткі множини в різних просторах.

Нехай $f-$ відображення з $X = \{x\}$ в $Y = \{y\}$, причому змінній $x$ відповідає вхідний вплив (причина), а змінній $y-$ відповідний вихід (наслідок). Припустимо, що цілі задані як нечіткі множини $\tilde{G}_1, \tilde{G}_2, ..., \tilde{G}_n$ в $Y$, у той час як обмеження – нечіткі множини $\tilde{C}_1, \tilde{C}_2, ..., \tilde{C}_m$ в просторі $X$. Маючи нечітку множину $\tilde{G}_i$ в $Y$, можна знайти нечітку множину $\overline{G}_i$ в $X$, яке індукує нечітку множину $\tilde{G}_i$ в $Y$. Функція належності $\overline{G}_i$ задається рівністю

$$\mu_{\overline{G}_i}(x) = \mu_{\tilde{G}_i}(f(x)), \quad i = \overline{1,n}. \tag{12.2}$$

Після цього рішення $\tilde{D}$ може бути виражене перетином множин $\overline{G}_1, \overline{G}_2, ..., \overline{G}_n$ і $\tilde{C}_1, \tilde{C}_2, ..., \tilde{C}_m$. Використовуючи співвідношення (12.2), можна записати $\mu_{\tilde{D}}(x)$ у розгорнутому вигляді:

$$\mu_{\tilde{D}}(x) = \mu_{\tilde{G}_1}(f(x)) \wedge ... \wedge \mu_{\tilde{G}_n}(f(x)) \wedge \mu_{\tilde{C}_1}(x) \wedge ... \wedge \mu_{\tilde{C}_m}(x),$$

де $f : X \to Y$.

Таким чином, випадок, коли цілі й обмеження задаються як нечіткі множини в різних просторах, може бути зведений до випадку, коли вони задаються в одному і тому ж просторі.

## 12.2 Нечіткий багатокритеріальний аналіз об'єктів захисту інформації

Один з підходів до призначення «ваг» кінцевому набору $n$ порівнюваних об'єктів на основі матриці парних порівнянь був

запропонований Т. Сааті[5] в 1970 році [16]. Згодом цей підхід оформився в цілий розділ прийняття рішень при наявності одного, а також декількох критеріїв і отримав назву методу аналізу ієрархій.

Метод аналізу ієрархій (МАІ) – математичний інструмент системного підходу до складних проблем прийняття рішень. Він не дає відповіді на питання, що правильно, а що ні, але дозволяє людині, що приймає рішення, оцінити, який з розглянутих ним варіантів найкраще задовольняє його потребам і його розумінню проблеми (задачі). У його основі поряд з математикою закладені і психологічні аспекти. МАІ дозволяє зрозумілим та раціональним чином структурувати складну проблему прийняття рішень у вигляді ієрархії, порівняти і виконати кількісну оцінку альтернативних варіантів рішення. В даний час МАІ міцно увійшов в теорію і практику багатокритеріального вибору.

Нехай $X = \{x_1, x_2, ..., x_n\}$ – множина об'єктів захисту інформації, які підлягають багатокритеріальному аналізу, а $C = \{C_1, C_2, ..., C_m\}$ – множина кількісних та якісних критеріїв, якими оцінюються об'єкти.

Завдання полягає в упорядкуванні елементів множини $X$ за критеріями з множини $C$. Вирішення цієї задачі складається з наступних етапів:

1) розгляд критеріїв як нечітких множин, які задані на універсальних множинах об'єктів захисту інформації за допомогою функції належності;

2) визначення функцій належності нечітких множин на основі експертної інформації про парні порівняння об'єктів захисту інформації за допомогою дев'ятибальної шкали Сааті;

3) ранжування об'єктів захисту інформації на основі перетину нечітких множин – критеріїв, які відповідають відомій в теорії прийняття рішень схемі Белмана-Заде;

4) ранжування критеріїв методом парних порівнянь та облік

[5] Т. Сааті (нар. в 1926 р.) – американський математик.

отриманих рангів як ступенів концентрації відповідних функції належності.

Без проведення оцінки інформації неможливо адекватно оцінити доцільність витрат грошей і ресурсів на її захист. Цінність інформації обов'язково повинна враховуватися при виборі захисних заходів.

При проведенні оцінки інформації було розглянуто три об’єкти захисту.

*Перший об’єкт захисту* містить у собі детальний опис технології виробництва певного виробу. Технологія – це промислові або інші процеси, які передбачають використання наукових або інших знань для вирішення певних проблем або випуску певної продукції. Технологія включає в себе методи, прийоми, режим роботи, послідовність операцій та процедур; вона тісно пов'язана з обладнанням, інструментами, застосовуваними засобами та використовуваними матеріалами. Отже, ця інформація дуже важлива як для розробки та вдосконалення конструкції виробу, так і для його використання в наукових дослідженнях та промисловості.

*Другий об’єкт захисту* містить дані про конструкцію виробу. У процесі розробки конструкції виробу були враховані особливості технології його виробництва. Це просте та ефективне рішення дозволило істотно скоротити терміни та вартість виробництва. Дані цього набору важливі, оскільки вони містять основні конструктивні особливості виробу, фізичні властивості та параметри елементів конструкції, які надалі й визначатимуть характеристики виробу, такі як його продуктивність та напрацювання до відмови.

*Третій об’єкт захисту* являє собою дані про функціональну схему виробу. Функціональна схема розкриває принцип дії функціонального блоку або вузла, тобто показує логіку роботи. Наявність функціональної схеми дозволяє бачити його принципові недоліки та переваги, а також значно спрощує його подальшу модифікацію, аналіз відмов та ремонт. Тому ця інформація також є дуже важливою, оскільки являє собою не просто дані, а чітко структурований набір даних, тобто такий, в якому переважає саме структурна складова.

Для оцінки інформації необхідно застосувати нечіткий багатокритеріальний аналіз інформації, що зберігається в об'єктах ЗІ по деяким найбільш важливим критеріям, а саме:

1) за об'ємом;
2) за новизною;
3) за перспективністю використання;
4) за ступенем завершеності;
5) за правовою захищеністю;
6) за практичною застосовністю.

Позначимо кожен з критеріїв – $C_i$, $i = \overline{1,6}$.

За першим критерієм $C_1$ - об'ємом інформації - технологічна документація (ТД) як комплекс графічних та текстових документів, що містять дані для організації виробничого процесу, має помірну перевагу над конструкторською документацією (КД), як комплексом графічних та текстових документів, які окремо або в сукупності визначають склад та налаштування виробу, оскільки за змістом ТД містить переважно словесні описи великого текстового об'єму, а КД – креслення у виді графічних документів. Відповідно, функціональна схема (ФС) містить ще менше інформації, ніж ТД, оскільки на ній зображають найбільш важливі блоки системи та зв'язки між ними. Тобто ТД має істотну перевагу над ФС за критерієм «об'єм».

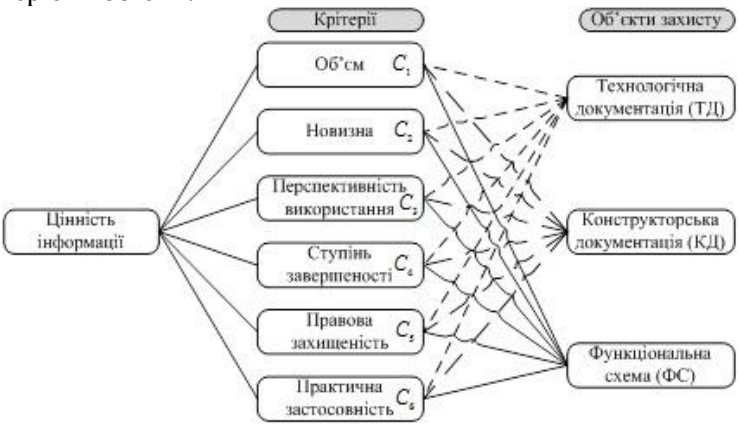



Рис. 12.1.Показники цінності інформації в об'єктах захисту інформації.

За другим критерієм $C_2$ - новизною - ТД та КД мають однаковий рівень, оскільки обидва документи розробляються вперше, тоді як до ФС була отримана внесенням деяких удосконалень у стару ФС, тому ТД має помірну перевагу над ФС.

За третім критерієм $C_3$ - перспективністю використання - ТД та ФС вносять однаковий внесок у досягнення мети, оскільки вони обидва дуже перспективні. КД трохи менш перспективна, тому ТД має помірну перевагу над КД.

За четвертим критерієм $C_4$ - ступенем завершеності - ФС має значну перевагу на ТД, оскільки ТД потребує ще значної доробки після того, як доробиться КД, яка має помірну перевагу над ТД.

За п'ятим критерієм $C_5$ - правовою захищеністю - КД вже захищена патентом, тоді як ФС взагалі не планується захищати патентом, а ТД, оскільки ще не дороблена – теж ще не захищена. Тому КД має істотну перевагу над ТД та значну перевагу над ФС.

За шостим критерієм $C_6$ - практичною застосовністю - ФС має найбільшу перевагу, оскільки досить внести незначне удосконалення, як можемо отримати новий виріб. Тому ФС має помірну перевагу над ТД та майже значну над КД.

## 12.3 Метод парних порівнянь об'єктів захисту інформації

Парні порівняння об'єктів захисту проводять за дев'ятибальною шкалою Сааті, представленою в таблиці 12.1.

Таблиця 12.1.

Шкала відношень Сааті

| Бал | Пояснення |
|---|---|
| 1 | Відсутня перевага об'єкта $x_j$ над об'єктом $x_i$ |
| 3 | Помірна перевага об'єкта $x_j$ над об'єктом $x_i$ |
| 5 | Істотна (сильна) перевага об'єкта $x_j$ над об'єктом $x_i$ |
| 7 | Значна (очевидна) перевага об'єкта $x_j$ над об'єктом $x_i$ |
| 9 | Абсолютна перевага об'єкта $x_j$ над об'єктом $x_i$ |
| 2, 4, 6, 8 | Проміжні порівняльні оцінки |

Порівняння наборів даних за критеріями $C_i$, $i=\overline{1,6}$ у відповідності до таблиці 1 приводить до наступних висловлювань:

$$\text{критерій } C_1: \begin{cases} \text{помірна перевага (3) } x_1 \text{ над } x_2, \\ \text{істотна перевага (5) } x_1 \text{ над } x_3; \end{cases}$$

$$\text{критерій } C_2: \begin{cases} \text{відсутня перевага (1) } x_1 \text{ над } x_2, \\ \text{помірна перевага (3) } x_1 \text{ над } x_3; \end{cases}$$

$$\text{критерій } C_3: \begin{cases} \text{помірна перевага (3) } x_1 \text{ над } x_2, \\ \text{відсутня перевага (1) } x_1 \text{ над } x_3; \end{cases}$$

$$\text{критерій } C_4: \begin{cases} \text{помірна перевага (3) } x_2 \text{ над } x_1, \\ \text{значна перевага (7) } x_3 \text{ над } x_1; \end{cases}$$

$$\text{критерій } C_5: \begin{cases} \text{істотна перевага (5) } x_2 \text{ над } x_1, \\ \text{значна перевага (7) } x_2 \text{ над } x_3; \end{cases}$$

$$\text{критерій } C_6: \begin{cases} \text{помірна перевага (3) } x_3 \text{ над } x_1, \\ \text{майже значна перевага (6) } x_3 \text{ над } x_2; \end{cases}$$

Далі необхідно сформувати матриці парних порівнянь об'єктів захисту інформації за кожним критерієм. Загальна кількість таких матриць співпадає з кількістю критеріїв і дорівнює $m$.

Нехай заданий набір з $n$ об'єктів (елементів), які позначимо $x_1, x_2, \ldots, x_n$. Для критерію $C_k$ матриця парних порівнянь має вигляд:

$$A(C_k) = \left(a_{ij}\right)_{n\times n} = \begin{array}{c} \\ x_1 \\ x_2 \\ \ldots \\ x_n \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & \ldots & x_n \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \end{array}, \quad i,j=\overline{1,n} \quad k=\overline{1,m}. \qquad (12.3)$$

де елемент $a_{ij}$ оцінюється експертом за дев'ятибальною шкалою

Сааті і являє собою відношення ваги об'єкта $C_i$ до ваги об'єкта $C_j$.

Знання матриці (12.3) дозволяє за допомогою методу Сааті проранжувати кожний об'єкт $x_i$ за кожним критерієм $C_k$. Для обчислення рангів необхідно знайти власний вектор матриці (12.3).

Матриця (12.3) має наступні властивості, які неважко перевірити безпосередньо:

1) Матриця діагональна, тобто всі розташовані на головній діагоналі елементи дорівнюють одиниці: $a_{ij} = 1, \quad i, j = \overline{1,n}, \quad i = j;$

2) Матриця зворотно-симетрична, тобто її елементи, що розташовані симетрично відносно головної діагоналі, є зворотними по відношенню один до одного: $a_{ij} = \dfrac{1}{a_{ji}}$ для всіх $i, j = \overline{1,n};$

3) Матриця транзитивна, тобто $a_{il} \cdot a_{lj} = a_{ij}$.

Наявність цих властивостей дозволяє визначити всі елементи матриці (12.3) за елементами одного з рядків. Якщо відомий $l-$й рядок, тобто елементи $a_{lj}$, $j = \overline{1,n}$, то довільний елемент $a_{ij}$ визначається наступним чином:

$$a_{ij} = \frac{a_{lj}}{a_{li}}, \quad i, j, l = \overline{1,n}.$$

*Приклад 1.*

$$A(C_1) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \end{array} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 5/3 \\ 1/5 & 3/5 & 1 \end{pmatrix} \end{array} \Rightarrow$$

$$a_{32} = \frac{a_{12}}{a_{13}} \Rightarrow 3/5 = \frac{3}{5}.$$

Експертним висловлюванням про об'єкти захисту відповідають такі матриці парних порівнянь:

$$A(C_1)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 5/3 \\ 1/5 & 3/5 & 1 \end{pmatrix} \end{array} \quad A(C_2)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{pmatrix} \end{array}$$

$$A(C_3)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 3 & 1 \\ 1/3 & 1 & 1/3 \\ 1 & 3 & 1 \end{pmatrix} \end{array} \quad A(C_4)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 1/3 & 3 \\ 3 & 1 & 3/7 \\ 7 & 7/3 & 1 \end{pmatrix} \end{array}$$

$$A(C_5)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 1/5 & 7/5 \\ 5 & 1 & 7 \\ 5/7 & 1/7 & 1 \end{pmatrix} \end{array} \quad A(C_6)=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 2 & 1/3 \\ 1/2 & 1 & 1/6 \\ 3 & 6 & 1 \end{pmatrix} \end{array}$$

Проведені розрахунки дають такі результати:

*C1 =*

| | | |
|---|---|---|
| *1.0000* | ***3.0000*** | ***5.0000*** |
| *0.3333* | *1.0000* | *1.6667* |
| *0.2000* | *0.6000* | *1.0000* |

*C2 =*

| | | |
|---|---|---|
| *1.0000* | ***1.0000*** | ***3.0000*** |
| *1.0000* | *1.0000* | *3.0000* |
| *0.3333* | *0.3333* | *1.0000* |

*C3 =*

| | | |
|---|---|---|
| *1.0000* | ***3.0000*** | ***1.0000*** |
| *0.3333* | *1.0000* | *0.3333* |
| *1.0000* | *3.0000* | *1.0000* |

*C4 =*

| | | |
|---|---|---|
| *1.0000* | *0.3333* | *0.1429* |
| ***3.0000*** | *1.0000* | *0.4286* |
| ***7.0000*** | *2.3333* | *1.0000* |

*C5 =*

| | | |
|---|---|---|
| *1.0000* | *0.2000* | *1.4000* |
| ***5.0000*** | *1.0000* | ***7.0000*** |
| *0.7143* | *0.1429* | *1.0000* |

*C6 =*

| | | |
|---|---|---|
| *1.0000* | *2.0000* | *0.3333* |
| *0.5000* | *1.0000* | *0.1667* |
| ***3.0000*** | ***6.0000*** | *1.0000* |

У цих матрицях виділені елементи відповідають парним порівнянням. Інші елементи знайдені в припущенні про

узгодженість парних порівнянь, тобто з урахуванням того, що матриця парних порівнянь є діагональною та має властивості транзитивності та зворотної симетричності.

Нехай $\mu_{C_k}(x_i)-$ число в діапазоні [*0, 1*], яке характеризує рівень оцінки об'єкту $x_i$ за критерієм $C_k$. Чим більше число $\mu_{C_k}(x_i),$ тим вище оцінка об'єкту $x_i$ за критерієм $C_k,\ \ i=\overline{1,n},\ \ k=\overline{1,m}.$. Тоді критерій $C_k$ можна представити у вигляді нечіткої множини $\tilde{C}_k$, яка задана на універсальній множині $X$ наступним чином:

$$\tilde{C}_k=\left\{\frac{\mu_{C_k}(x_1)}{x_1},\frac{\mu_{C_k}(x_2)}{x_2},...,\frac{\mu_{C_k}(x_n)}{x_n}\right\},$$

де $\mu_{C_k}(x_i)-$ ступінь належності елемента $x_i$ до нечіткої множини $\tilde{C}_k$.

Після визначення всіх елементів матриці парних порівнянь (12.3) ступені належності, необхідні для формування нечіткої множини, обчислюються за формулою:

$$\mu_{C_k}(x_i)=\frac{1}{a_{1i}+a_{2i}+...+a_{ni}}. \qquad (12.4)$$

Користуючись матрицями парних порівнянь і виразом (12.4), отримуємо:

*G1 = 0.6522 0.2174 0.1304G2 = 0.6522 0.4286 0.1429*
*G3 = 0.4286 0.1429 0.4286G4 = 0.0909 0.2727 0.6364*
*G5 = 0.1489 0.7447 0.1064G6 = 0.2222 0.1111 0.6667*

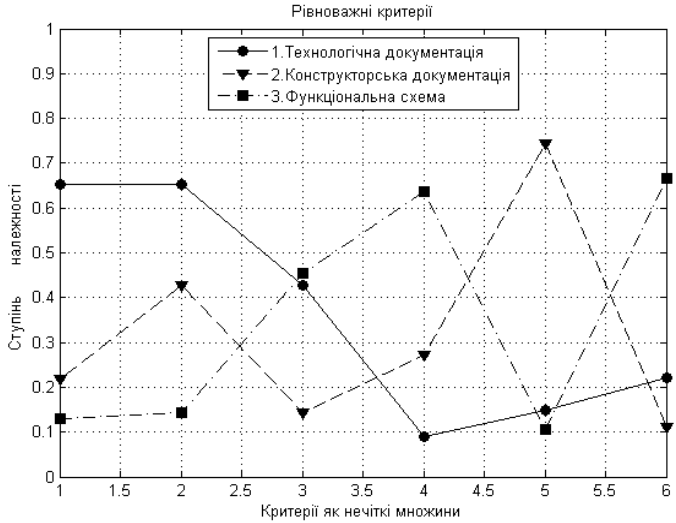



Рис 12.2. Об'єкти захисту на нечітких множинах рівноважних критеріїв.

Більш докладно процес обчислення виглядає наступним чином (для прикладу обрахуємо ступені належності $x_i$, $i=\overline{1,n}$ до нечіткої множини $\tilde{C}_1$):

$$\mu_{C_1}(x_1)=\frac{1}{a_{11}+a_{21}+a_{31}}\Rightarrow\frac{1}{1+1/3+1/5}=\frac{1}{1.00+0.33+0.20}=$$

$$=\frac{1}{1.53}=0.652$$

$$\mu_{C_1}(x_2)=\frac{1}{a_{12}+a_{22}+a_{32}}\Rightarrow\frac{1}{3+1+3/5}=\frac{1}{3.00+1.00+0.60}=$$

$$=\frac{1}{5.67}=0.2174$$

$$\mu_{C_1}(x_3) = \frac{1}{a_{13} + a_{23} + a_{33}} \Rightarrow \frac{1}{5 + 5/3 + 1} = \frac{1}{1.00 + 0.33 + 1.00} =$$

$$= \frac{1}{7.67} = 0.1304..$$

$$\tilde{C}_1 = \left\{ \frac{0.6522}{x_1}, \frac{0.2174}{x_2}, \frac{0.1304}{x_3} \right\}, \tilde{C}_2 = \left\{ \frac{0.6522}{x_1}, \frac{0.4286}{x_2}, \frac{0.1429}{x_3} \right\},$$

$$\tilde{C}_3 = \left\{ \frac{0.4286}{x_1}, \frac{0.1429}{x_2}, \frac{0.4286}{x_3} \right\}, \tilde{C}_4 = \left\{ \frac{0.0909}{x_1}, \frac{0.2727}{x_2}, \frac{0.6364}{x_3} \right\},$$

$$\tilde{C}_5 = \left\{ \frac{0.1489}{x_1}, \frac{0.7447}{x_2}, \frac{0.1064}{x_3} \right\}, \tilde{C}_6 = \left\{ \frac{0.2222}{x_1}, \frac{0.1111}{x_2}, \frac{0.6667}{x_3} \right\}.$$

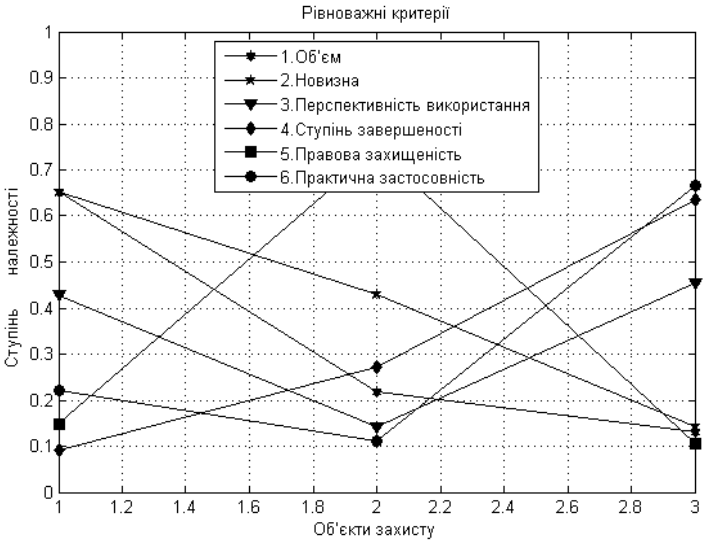



Рис. 12.3. Рівноважні критерії як нечіткі множини.

## 12.4 Рівноважні і нерівноважні критерії

Базуючись на принципі Белмана-Заде [20], найкращою системою будемо вважати ту, яка одночасно краща за всіма

критеріями. Тому нечітка множина, яка необхідна для рейтингового аналізу, визначається у вигляді перетину:

$$\tilde{D} = \tilde{C}_1 \cap \tilde{C}_2 \cap ... \cap \tilde{C}_m.$$

Враховуючи той факт, що в теорії нечітких множин операції перетину $\cap$ відповідає $min$, отримуємо:

$$\tilde{D} = \left\{ \frac{\min\limits_{k=1,m} \mu_{C_k}(x_1)}{x_1}, \frac{\min\limits_{k=1,m} \mu_{C_k}(x_2)}{x_2}, ..., \frac{\min\limits_{k=1,m} \mu_{C_k}(x_n)}{x_n} \right\} \qquad (12.5)$$

Користуючись нечіткими множинами $\tilde{C}_1, \tilde{C}_2, ..., \tilde{C}_6$ та виразом (12.5), знаходимо:

$$D = 0.0909 \quad 0.1111 \quad 0.1064$$

$$\tilde{D} = \left\{ \frac{0.0909}{x_1}, \frac{0.1111}{x_2}, \frac{0.1064}{x_3} \right\}.$$

Після нормування:

$$D_{norm} = 0.2948 \quad 0.3603 \quad 0.3449$$

$$\tilde{D}_{norm} = \left\{ \frac{0.2948}{x_1}, \frac{0.3603}{x_2}, \frac{0.3449}{x_3} \right\}.$$

Результати свідчать про перевагу цінності інформації на об'єкті $x_2$ (КД) над цінністю інформації на об'єктах $x_1$ (ТД) та $x_3$ (ФС), а також про перевагу цінності інформації на об'єкті $x_3$ (ФС) над цінністю інформації на об'єкті $x_1$ (ТД).

Нехай $w_1, w_2, ... w_m -$ коефіцієнти відносної важливості (або ранги) критеріїв $C_1, C_2, ..., C_m$, такі що $w_1 + w_2 + ... + w_m = 1$. Для визначення коефіцієнтів $w_j$, $j = \overline{1,m}$ необхідно сформувати матрицю парних порівнянь важливості критеріїв $C_k$, аналогічну матриці (12.3) і скористатися формулою (12.4).

Найважливішими критеріями вважаються «Перспективність використання» та «Практична застосовність». Не менш важливими

вважається критерій «Правова захищеність», оскільки патентування має такі переваги:

- отримання комерційної вигоди від безперешкодного випуску запатентованого виробу;
- патент можна продати або надати право на його використання;
- патент є нематеріальним активом, вартість якого можна внести в статутний капітал підприємства;
- наявність патенту часто служить рекламним цілям.

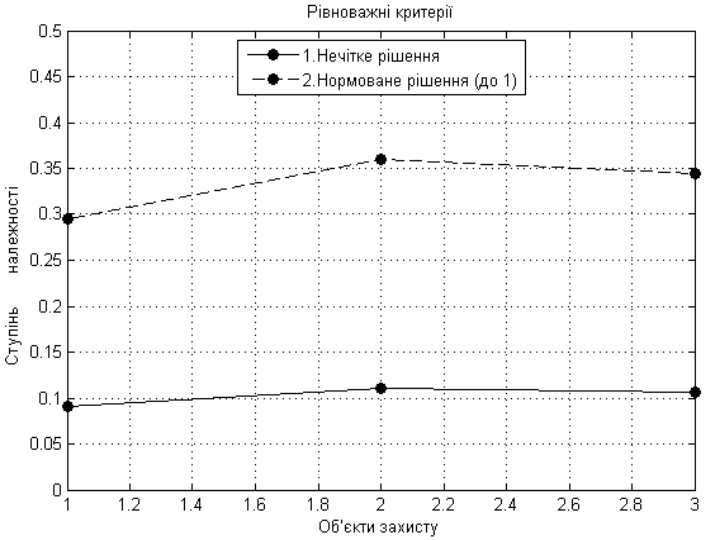



Рис. 12.4. Нечітке та нормоване до одиниці рішення при рівноважних критеріях.

Критерій «Новизна» є також важливим, оскільки характеризує виявлення раніше невідомих властивостей, явищ, закономірностей, зв'язків, співвідношень, методів, схем, форм, параметрів або процесів.

Таким чином, можна записати наступні висловлювання:

ранги критеріїв $\begin{cases} \text{істотна перевага (5) } C_2 \text{ над } C_1 \\ \text{майже абсолютна перевага (8) } C_3 \text{ над } C_1 \\ \text{майже значна перевага (6) } C_4 \text{ над } C_1 \\ \text{значна перевага (7) } C_5 \text{ над } C_1 \\ \text{майже абсолютна перевага (8) } C_6 \text{ над } C_1 \end{cases}$

Експертними оцінками відповідає наступна матриця парних порівнянь:

$$A = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{array} \begin{array}{c} \begin{array}{cccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 \end{array} \\ \begin{pmatrix} 1 & 1/5 & 1/8 & 1/6 & 1/7 & 1/8 \\ 5 & 1 & 5/8 & 5/6 & 5/7 & 5/8 \\ 8 & 8/5 & 1 & 8/6 & 8/7 & 8/8 \\ 6 & 6/5 & 6/8 & 1 & 6/7 & 6/8 \\ 7 & 7/5 & 7/8 & 7/6 & 1 & 7/8 \\ 8 & 8/5 & 8/8 & 8/6 & 8/7 & 1 \end{pmatrix} \end{array}$$

В результаті розрахунків маємо:

$A =$

| | | | | | |
|---|---|---|---|---|---|
| *1.0000* | *0.2000* | *0.1250* | *0.1667* | *0.1429* | *0.1250* |
| ***5.0000*** | *1.0000* | *0.6250* | *0.8333* | *0.7143* | *0.6250* |
| ***8.0000*** | *1.6000* | *1.0000* | *1.3333* | *1.1429* | *1.0000* |
| ***6.0000*** | *1.2000* | *0.7500* | *1.0000* | *0.8571* | *0.7500* |
| ***7.0000*** | *1.4000* | *0.8750* | *1.1667* | *1.0000* | *0.8750* |
| ***8.0000*** | *1.6000* | *1.0000* | *1.6000* | *1.1429* | *1.0000* |

Для знаходження власного вектора, а отже й рангів критеріїв $C_1, C_2, ..., C_6$, необхідно вирішити характеристичне рівняння шостого степеню $(A - \lambda \cdot E) \cdot w = 0$, де $E -$ одинична матриця, $w -$ власний вектор. За допомогою програми MATLABця задача вирішується просто. Найбільший власний вектор, знайдений таким чином, визначає ранги критеріїв:

*Vmax = 0.0641 0.3205 0.5129 0.3847 0.4488 0.5299*

$\hat{w}_1 = 0.0641$, $\hat{w}_2 = 0.3205$, $\hat{w}_3 = 0.5129$, $\hat{w}_4 = 0.3847$, $\hat{w}_5 = 0.4488$, $\hat{w}_6 = 0.5299$.

Після нормування ранги критеріїв приймають значення:

*Vnorm = 0.0284 0.1418 0.2269 0.1701 0.1985 0.2344*

$w_1 = 0.0284$, $w_2 = 0.1418$, $w_3 = 0.2269$, $w_4 = 0.1701$, $w_5 = 0.1985$, $w_6 = 0.2344$.

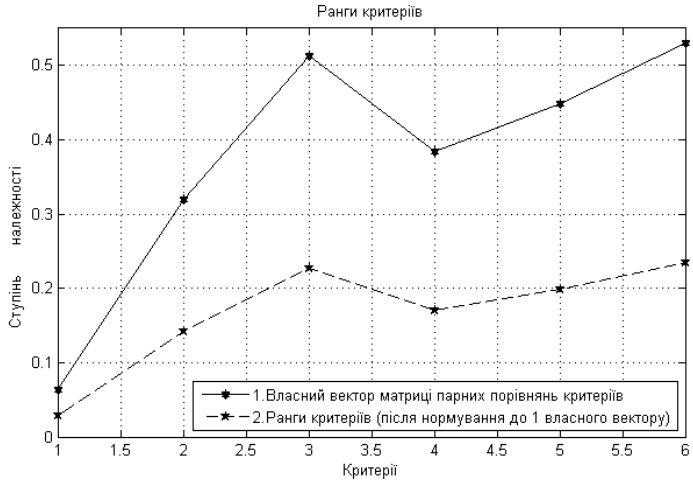



Рис.12.5. Власний вектор матриці парних порівнянь критеріїв та їх ранги.

При наявності рангів $w_j$, $j = \overline{1,m}$ формула (5) прийме вигляд

$$\tilde{D} = \left\{ \frac{\underline{min}_{k=1,m}\left(\mu_{C_k}(x_1)\right)^{w_1}}{x_1}, \frac{\underline{min}_{k=1,m}\left(\mu_{C_k}(x_2)\right)^{w_2}}{x_2}, ..., \frac{\underline{min}_{k=1,m}\left(\mu_{C_k}(x_n)\right)^{w_m}}{x_n} \right\}, \quad (12.6)$$

де $w_j$ свідчить про концентрацію нечіткої множини $\tilde{C}_k$ відповідно до міри важливості критерію $C_k$. Тоді згідно (6) одержуємо:

$$
\begin{aligned}
Gn1 &= 0.9880 \quad 0.9576 \quad 0.9439 \\
Gn2 &= 0.9412 \quad 0.8868 \quad 0.7589 \\
Gn3 &= 0.8251 \quad 0.6431 \quad 0.8251 \\
Gn4 &= 0.6650 \quad 0.8017 \quad 0.9260 \\
Gn5 &= 0.6852 \quad 0.9432 \quad 0.6410 \\
Gn6 &= 0.7029 \quad 0.5975 \quad 0.9093
\end{aligned}
$$

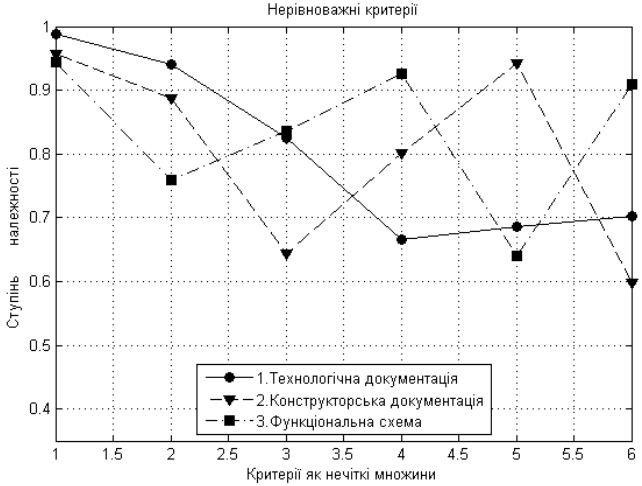



Рис. 12.6.Об’єкти захисту нанечітких множинах нерівноважних критеріїв.

$$
\tilde{C}_1 = \left\{ \frac{0.6522^{0.0284}}{x_1}, \frac{0.2174^{0.0284}}{x_2}, \frac{0.1304^{0.0284}}{x_3} \right\} =
$$

$$
= \left\{ \frac{0.9880}{x_1}, \frac{0.9576}{x_2}, \frac{09439}{x_3} \right\},
$$

$$
\tilde{C}_2 = \left\{ \frac{0.6522^{0.1418}}{x_1}, \frac{0.4286^{0.1418}}{x_2}, \frac{0.1429^{0.1418}}{x_3} \right\} =
$$

$$
= \left\{ \frac{0.9412}{x_1}, \frac{0.8868}{x_2}, \frac{0.7589}{x_3} \right\},
$$

$$\tilde{C}_3 = \left\{ \frac{0.4286^{0.2269}}{x_1}, \frac{0.1429^{0.2269}}{x_2}, \frac{0.4286^{0.2269}}{x_3} \right\} =$$

$$= \left\{ \frac{0.8251}{x_1}, \frac{0.6431}{x_2}, \frac{0.8251}{x_3} \right\},$$

$$\tilde{C}_4 = \left\{ \frac{0.0909^{0.1701}}{x_1}, \frac{0.2727^{0.1701}}{x_2}, \frac{0.6364^{0.1701}}{x_3} \right\} =$$

$$= \left\{ \frac{0.6650}{x_1}, \frac{0.8017}{x_2}, \frac{0.9260}{x_3} \right\},$$

$$\tilde{C}_5 = \left\{ \frac{0.1489^{0.1985}}{x_1}, \frac{0.7447^{0.1985}}{x_2}, \frac{0.1064^{0.1985}}{x_3} \right\} =$$

$$= \left\{ \frac{0.6852}{x_1}, \frac{0.9432}{x_2}, \frac{0.6410}{x_3} \right\},$$

$$\tilde{C}_6 = \left\{ \frac{0.2222^{0.2344}}{x_1}, \frac{0.1111^{0.2344}}{x_2}, \frac{0.6667^{0.2344}}{x_3} \right\} =$$

$$= \left\{ \frac{0.7029}{x_1}, \frac{0.5975}{x_2}, \frac{0.9093}{x_3} \right\}.$$

Перетин цих нечітких множин з урахуванням рангів критеріїв має вигляд:

*Dn = 0.6650 0.5975 0.6410*

$$\tilde{D} = \left\{ \frac{0.6650}{x_1}, \frac{0.5975}{x_2}, \frac{0.6410}{x_3} \right\}$$

Після нормування нечітке рішення приймає такі значення:

*Dnorm = 0.3494 0.3139 0.3367*

$$\tilde{D}_{norm} = \left\{ \frac{0.3494}{x_1}, \frac{0.3139}{x_2}, \frac{0.3367}{x_3} \right\}$$

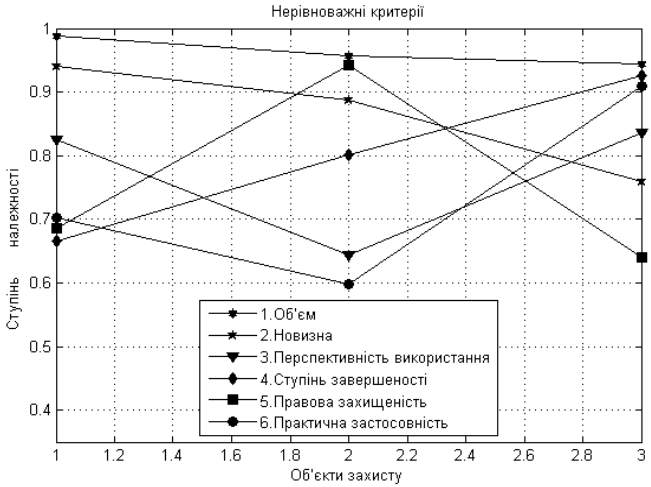



Рис. 12.7. Нерівноважні критерії як нечіткі множини.

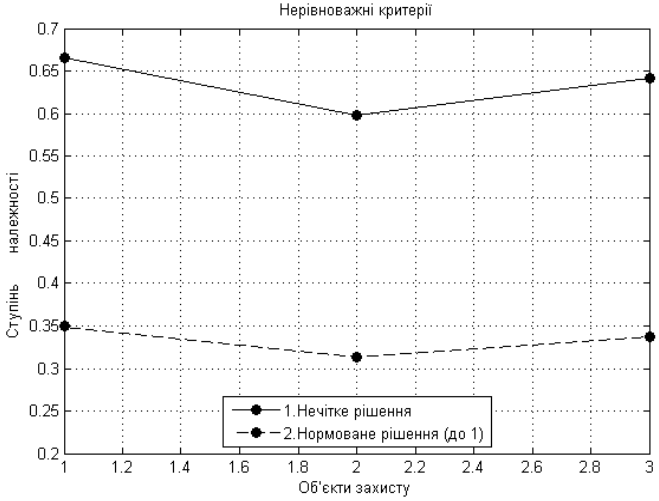



Рис. 12.8.Нечітке та нормоване до 1 рішення при нерівноважних критеріях.

Це свідчить про значну перевагу цінності інформації на об’єкті $x_1$ над цінністю інформації на об’єктах $x_2$ і $x_3$, а також про

помірну перевагу цінності інформації на об'єкті $x_3$ над цінністю інформації на об'єкті $x_2$.

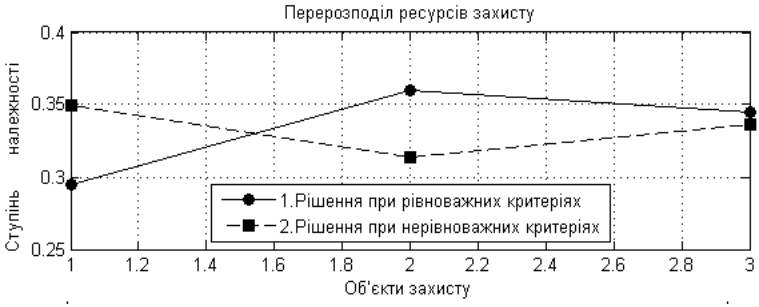



Рис. 12.9. Перерозподіл ресурсів захисту.

Після врахування рангів критеріїв виявилось, що розподіл ресурсів між об'єктами захисту кардинально змінюється, тобто об'єкт захисту «Технологічна документація» потребує більшої уваги, ніж об'єкт захисту «Конструкторська документація», на відміну від випадку, коли критерії вважаються рівноважними.

## 12.5 Матричний підхід до багатокритеріального аналізу ризиків інформаційної безпеки

Аналіз ризиків є важливим завданням менеджменту інформаційної безпеки. Існує декілька різноманітних методик та методологій, розробленими провідними вченими у цій галузі. В даній статті розглядається методологія OCTAVE (Operationally Critical Threat, Assetand Vulnerability Evaluation – дослівно «Операційна оцінка критичних загроз, активів та вразливостей»), розроблена в Інституті програмної інженерії при Університеті Карнегі-Меллона та вдосконалена професором Університету Олбані Санжеєм Гоелом за допомогою матричного підходу, який ми використовуватимемо надалі.

Методологія по черзі співвідносить між собою критичні активи, вразливості, загрози та засоби протидії та контролю в трьох матрицях. Для оцінки ступеня впливу введено чотири градації: «0» – відсутність впливу; «2» – слабкий вплив; «4» – помірний вплив; «8» – сильний вплив.

Перша матриця (табл. 12.2 ) характеризує вплив вразливостей елементів системи захисту на окремі критичні активи. Сукупний вплив вразливостей $D_j$ визначається за формулою:

$$D_j = \sum_{i=1}^{n} g_{ij} \cdot B_i,$$

де $j = \overline{1,m}$ – номер вразливості, $i = \overline{1,n}$ – номер критичного активу, $B_i$ – його відносна вартість, $g_{ij}$ – приведене в таблиці число, що характеризує ступінь впливу $j$ – тої вразливості на $i$ – тий критичний актив. Вразливості елементів системи захисту ранжовані за ступенем сукупного впливу $D_j$ (попередньо за ступенем відносної вартості $B_i$ ранжовані критичні активи).

Таблиця 12.2. Матриця вразливостей.

| $g_{ij}$ $i=\overline{1,n}$ $j=\overline{1,m}$ **Основні активи та цінності** | Репутація та довіра клієнтів | Конфіденційні секрети клієнтів | Втрачений збут | Цілісність інформації | Доступність сервісів та послуг | Телекомунікації та зв'язок | Програмне забезпечення | Апаратне забезпечення | *Сукупний вплив вразливостей* | *Відносний вплив вразливостей* |
|---|---|---|---|---|---|---|---|---|---|---|
| *Відносна вартість активів* $B_i$ | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | | $D_j$ |
| **Вразливості** | | | | | | | | | | |
| Компоненти захисту – міжмережеві екрани | 8 | 8 | 4 | 8 | 4 | 8 | 8 | 8 | 248 | 8 |
| Компоненти мережі – комутатори, маршрутизатори | 8 | 8 | 4 | 4 | 8 | 4 | 4 | 8 | 224 | 7 |
| Помилки конфігурації Екстранет-серверів | 8 | 8 | 2 | 4 | 8 | 4 | 8 | 2 | 214 | 6 |
| Системи фізичної безпеки та виявлення вторгнень | 4 | 8 | 8 | 4 | 4 | 4 | 2 | 8 | 196 | 5 |
| Архітектура апаратного забезпечення | 8 | 4 | 4 | 4 | 4 | 2 | 8 | 8 | 182 | 4 |
| Помилки конфігурації Інтранет-серверів | 2 | 8 | 2 | 4 | 8 | 4 | 8 | 2 | 166 | 3 |
| Робочі станції користувачів та ноутбуки | 4 | 8 | 2 | 4 | 4 | 1 | 4 | 8 | 158 | 2 |
| Відключення електроенергії | 2 | 0 | 4 | 2 | 8 | 4 | 2 | 2 | 100 | 1 |

Всі оцінки суб'єктивні і були отримані на основі експертної оцінки.

Так, наприклад актив «Цілісність інформації» за оцінками експертів сильно залежить від справності міжмережевих екранів, тоді як «Конфіденційні секрети клієнтів» зовсім не пов'язані з «Відключенням електроенергії».

Таблиця 12.3. Матриця загроз.

| $h_{jk}$ $j=\overline{1,m}$ $k=\overline{1,p}$ **Вразливості** | Компоненти захисту – міжмережеві екрани | Компоненти мережі – комутатори, маршрутизатори | Помилки конфігурації Екстранет-серверів | Системи фізичної безпеки та виявлення вторгнень | Архітектура апаратного забезпечення | Помилки конфігурації Інтранет-серверів | Робочі станції користувачів та ноутбуки | Відключення електроенергії | *Сукупна значущість загроз* | *Відносна значущість загроз* |
|---|---|---|---|---|---|---|---|---|---|---|
| *Відносний вплив вразливостей* $D_j$ | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | | $U_k$ |
| **Загрози** | | | | | | | | | | |
| НСД до інформації, системи та мережі | 8 | 8 | 8 | 8 | 8 | 2 | 8 | 2 | 264 | 8 |
| Серверний збій | 8 | 8 | 2 | 8 | 4 | 8 | 2 | 8 | 224 | 7 |
| Відмова в обслуговуванні | 8 | 8 | 2 | 2 | 8 | 4 | 2 | 2 | 192 | 6 |
| Шантаж та вимагання | 2 | 8 | 4 | 8 | 4 | 4 | 8 | 2 | 182 | 5 |
| Внутрішні зловмисні атаки (шахрайство та саботаж) | 4 | 2 | 4 | 8 | 4 | 2 | 8 | 2 | 150 | 4 |
| Помилки та недбалість персоналу | 4 | 4 | 4 | 4 | 4 | 2 | 4 | 2 | 136 | 3 |
| IP-спуфінг та маскування | 2 | 8 | 2 | 2 | 4 | 2 | 8 | 2 | 134 | 2 |
| Віруси або інші шкідливі програми | 2 | 4 | 2 | 2 | 4 | 2 | 8 | 2 | 106 | 1 |

В другій матриці (табл. 12.3) ми за такою ж методикою враховуємо можливості реалізації окремих типів загроз через вразливості елементів системи захисту. Сукупна значущість загроз $U_k$ визначається за формулою:

$$U_k = \sum_{j=1}^{m} h_{jk} \cdot D_j ,$$

де $k=\overline{1,p}-$ номер загрози, $h_{jk}-$ ступінь можливості реалізації $k-$ї загрози через $j-$ту вразливість елементу системи захисту. Ранжування ведеться за ступенем сукупної значущості загроз $U_k$.

Наприклад, реалізація загрози «Віруси або інші шкідливі програми» найбільш ймовірна через вразливість системи «Робочі станції користувачів та ноутбуки», оскільки занесення вірусу в мережу співробітниками дуже часто відбувається при перегляді ними електронних листів від невідомих людей або інфікованих веб-сторінок в Інтернеті.

Таблиця 12.4. Матриця контрзаходів.

| $s_{kl}$ $k=\overline{1,p}$ $l=\overline{1,q}$ **Загрози** | Проникнення (злом, атака на пароль) | Серверний збій | Відмова в обслуговуванні | Шантаж та здирство шляхом насильства | Внутрішні зловмисні атаки та шпигунство | Помилки та недбалість персоналу | IP-спуфінг та маскування | Шкідливі програми (віруси, черв'яки і т. д.) | *Сукупна ефективність контрзаходів* | *Відносна ефективність з контрзаходів* |
|---|---|---|---|---|---|---|---|---|---|---|
| *Відносна значущість загроз* $U_k$ | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | | $A_l$ |
| **Засоби протидії та контролю** | | | | | | | | | | |
| Політика безпеки | 8 | 2 | 4 | 8 | 8 | 4 | 8 | 8 | 210 | 7 |
| Міжмережеві захисні екрани | 8 | 8 | 2 | 4 | 2 | 4 | 8 | 2 | 190 | 6 |
| Фізичний захист навколишнього середовища | 8 | 4 | 4 | 2 | 8 | 4 | 2 | 4 | 178 | 5 |
| Навчання персоналу | 8 | 2 | 2 | 4 | 4 | 8 | 4 | 8 | 166 | 4 |
| Аудит та моніторинг – система виявлення вторгнень | 4 | 8 | 2 | 2 | 4 | 4 | 8 | 4 | 158 | 3 |
| Конфігурація архітектури | 2 | 8 | 8 | 2 | 4 | 2 | 2 | 0 | 156 | 2 |
| Демілітаризована зона | 4 | 8 | 2 | 4 | 0 | 4 | 4 | 0 | 140 | 1 |

В третій матриці (табл.12.4) ми враховуємо протидію окремих елементів захисту приведеним загрозам. Сукупна ефективність контрзаходів $A_l$ визначається за формулою:

$$A_l = \sum_{l=1}^{q} s_{kl} \cdot U_k ,$$

де $l = \overline{1,q}$ – номер контрзаходу, $s_{kl}$ – вплив $l$ – того контрзаходу на $k$ – ту загрозу.

Наприклад, саме контрзахід «Навчання персоналу» буде найбільш ефективним при усуванні загрози «Помилки та недбалість персоналу», оскільки професійні недоліки, відступ від робочих правил або порушення правил захисту занадто дорого обходяться фірмі.

Ранжування за цим принципом контрзаходів дозволяє в значній мірі оптимізувати СЗІ.

Наведену просту у виконанні та використанні методику легко застосувати для внутрішнього аудиту на відміну від дорогих методик аудиторських фірм, які лякають своєю громіздкістю та складністю.

## 12.6 Визначення інтервалів допустимих витрат на захист інформації

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 [20], в якій процес прийняття рішень ведеться в умовах невизначеності. При цьому і шукана величина, і обмеження, які сукупно формують цільову функцію, задаються нечіткими множинами.

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:

$Y = \{y\}$- 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:

$$\mu_G(y) = \frac{a}{a + b*(y - y_0)^2} \qquad (12.7)$$

Значення параметрів в (12.7) 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 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 дрібно-лінійних або дрібно-нелінійних функцій [13]. На першому етапі будемо використовувати дрібно-лінійні функції:

$$f(x,y) = \frac{x/y}{x/y + c} = \frac{1}{1 + c(y/x)} \qquad (12.8)$$

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

$$\mathrm{f(y)} = \frac{1}{1 + \tilde{c}y}, \text{ де } \tilde{c} = \frac{c}{x} . \qquad (12.9)$$

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

$$\mu_C(y) = \frac{cy}{1 + cy} . \qquad (12.10)$$

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), при цьому $\mu_C(y)$ зростає (при $y \to \infty$ до 1).

Графічно це виглядає так:

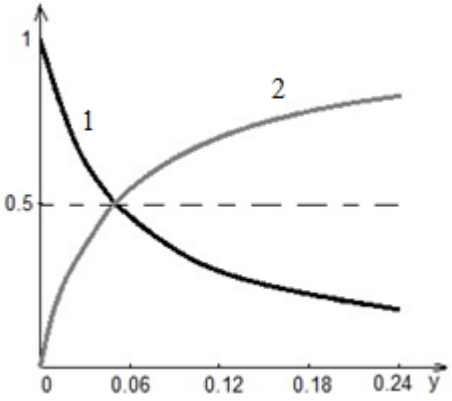



Рис. 12.10. Залежності від *y* частки вилученої інформації *f(x,y)* (крива 1) і функції належності $\mu_c(y)$ до нечіткої множини *C(y)* (крива 2).

Параметри *a, b, c*, які входять в (12.7), (12.8), визначаємо з наступних міркувань. Параметр *a* в (12.7) 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 (12.8) також визначається з суб'єктивних міркувань, а саме – заданим рівнем строгості виконання нечіткого обмеження.

Нашою метою є встановлення інтервалу допустимих значень ресурсів захисту, виділених на кожний об'єкт. Інтервал визначається по заданому рівню певного результуючого показника, який враховує ступінь виконання обох нечітких умов. Цей рівень задається менеджментом і залежить від наявної 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 (ступінь їх впливу і підлягає дослідженню) є:

1) форма функції $\mu_G(y)$, зокрема її ширина (параметр *b* в (12.7) ) 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 (12.7) );

3) кривизна функції $\mu_C(y)$ (параметр *c* в (12.8) ).

Вплив цих факторів видно з рис.13.10. 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).

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

$$f_1(x,y)=\frac{x/y}{x/y+16} \qquad f_2(x,y)=\frac{x/y}{x/y+10} \qquad (12.11)$$

які відображають різний рівень вразливості об'єктів ($f_2 > f_1$). Величини $c_1$=16, $c_2$=10 в цих виразах вибрані такими, що при $x/y=1$ дають значення $f_1 = 0{,}058$, $f_2 = 0{,}090$, які, на нашу думку, можуть відображати реальні ситуації. Задамо для прикладу значення $f_0$ = 0,1 (допустима частка втраченої інформації – 10%) і з (12.11) одержимо: для першого об'єкта $x/y=1{,}78$ і при обраному значенні *x=0,2* (ресурс нападу складає 20% від вартості інформації на об'єкті), маємо *y = 0,112* і з (10) $\mu_C^{(1)}=0{,}66$, для другого об'єкта, відповідно, $x/y=1{,}11$; *y*= 0,18 і з *(10)* $\mu_C^{(2)}=0{,}64$ (точки *A* і *B* на рис.12.11,а) ).

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

В кожному з приведених варіантів змінюється порівняно з попередніми один з типів функцій - $\mu_G$ або $\mu_C$. Це дає змогу оцінити вплив параметрів*b i c*, які входять в ці функції, на кінцевий результат.

На рис.12.11,а – г функції $\mu_G(y)$ мають симетричний характер, а $\mu_C(y)$ - дрібно-лінійний, на рис. 12.11,д $\mu_C(y)$ мають дрібно-нелінійний характер.

Результати розрахунків по *варіантах 1 - 5* приведені нижче.

*Варіант 1.*

Для першого об'єкта:

$$\mu_G{}^{(1)}(y)=\frac{0.08}{0.08+100(y-0.08)^2} \qquad \mu_C{}^{(1)}(y)=\frac{16y}{1+16y}$$

Для другого об'єкта:

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+100(y-0.11)^2} \qquad \mu_C{}^{(2)}(y)=\frac{10y}{1+10y}$$

*Варіант 2.*

$$\mu_G{}^{(1)}(y)=\frac{0.08}{0.08+400(y-0.08)^2} \qquad \mu_C{}^{(1)}(y)=\frac{16y}{1+16y}$$

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+400(y-0.11)^2} \qquad \mu_C{}^{(2)}(y)=\frac{10y}{1+10y}$$

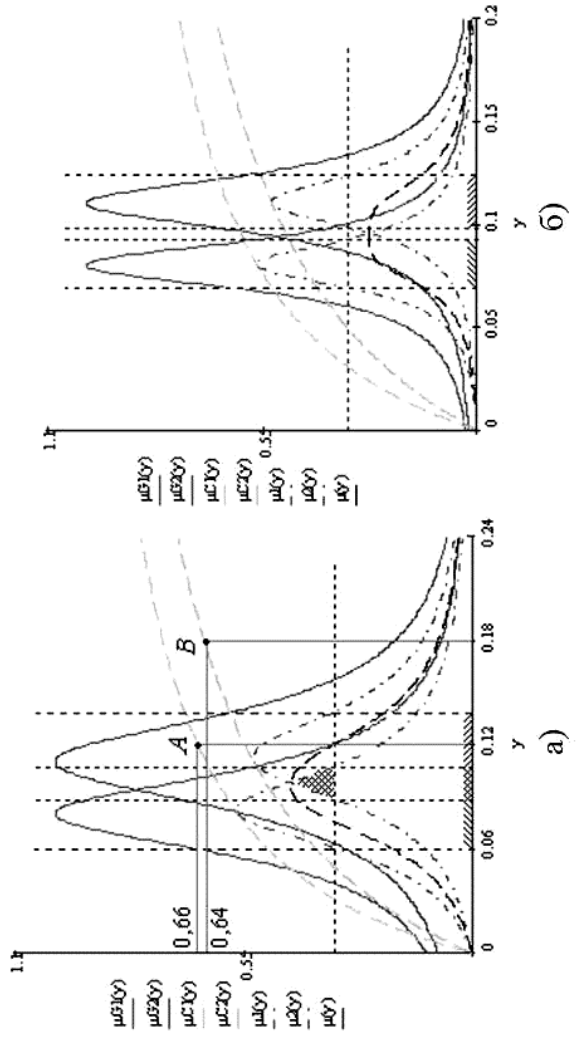


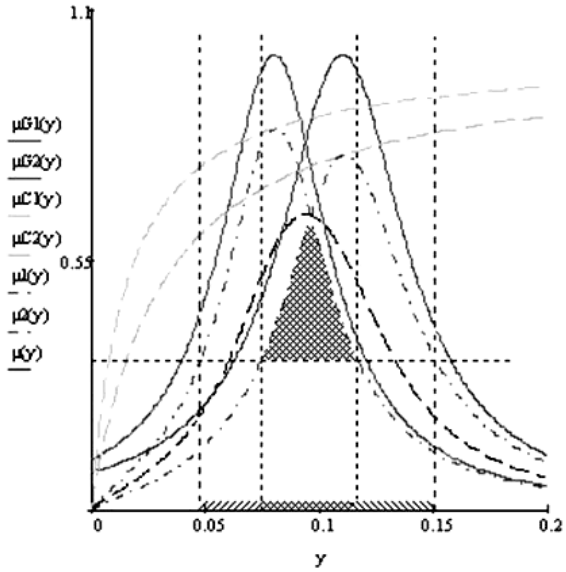

Рис. 12.11.

Функції належності $\mu_G$ і $\mu_C$ при різних значеннях параметрів *b і c*.

*Варіант 3*.

$$\mu_G{}^{(1)}(y)=\frac{0.08}{0.08+100(y-0.08)^2}\quad \mu_C{}^{(1)}(y)=\frac{64y}{1+64y}$$

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+100(y-0.11)^2}\quad \mu_C{}^{(2)}(y)=\frac{32y}{1+32y}$$




в)

Рис. 12.11. Функції належності $\mu_G$ і $\mu_C$ при різних значеннях параметрів *b i c.*

*Варіант 4.*

$$\mu_G{}^{(1)}(y)=\frac{0.08}{0.08+400(y-0.08)^2}\quad \mu_C{}^{(1)}(y)=\frac{64y}{1+64y}$$

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+400(y-0.11)^2}\quad \mu_C{}^{(2)}(y)=\frac{32y}{1+32y}$$

*Варіант 5.*

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

$$\mu_G{}^{(2)}(y)=\frac{0.11}{0.11+100(y-0.11)^2}\quad \mu_C{}^{(2)}(y)=\frac{90y^2}{1+90y^2}$$

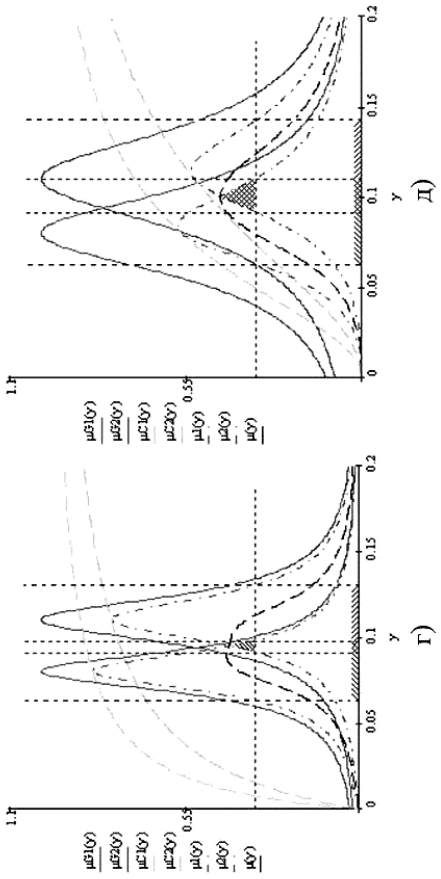



Рис. 12.11.

Функції належності $\mu_G$ і $\mu_C$ при різних значеннях параметрів *b і c.*

Результати приведених і подальших розрахунків зведені в табл. 12.5 , де введені позначення: $\Delta y^{(k)} = y_2^{(k)} - y_1^{(k)}$ - інтервал допустимих значень y, які задовольняють умові

$$\mu = \sqrt{\mu_G{}^{(k)}(y) \cdot \mu_C{}^{(k)}(y)} \geq 0{,}33;$$

$\Delta y$ - інтервал перекриття.

Інтервали $\Delta y^{(k)}$ позначені правонаправленою штриховкою для першого об'єкта і лівонаправленою – для другого.

Таблиця 12.5.

Допустимі інтервали ресурсів захисту

| № варіанта | Об'єкт №1 | | | | Об'єкт №2 | | | | Інтервал перекриття |
|---|---|---|---|---|---|---|---|---|---|
| | $y_1^{(1)}$ | $y_2^{(1)}$ | $\Delta y^{(1)}$ | $\Delta y^{(1)} / y_0^{(1)}$ | $y_1^{(2)}$ | $y_2^{(2)}$ | $\Delta y^{(2)}$ | $\Delta y^{(2)} / y_0^{(2)}$ | $\Delta y$ |
| 1 | 0.060 | 0.106 | 0.046 | 0.575 | 0.088 | 0.138 | 0.050 | 0.454 | 0.018 |
| 2 | 0.069 | 0.092 | 0.023 | 0.287 | 0.098 | 0.123 | 0.025 | 0.227 | 0 |
| 3 | 0.047 | 0.116 | 0.069 | 0.862 | 0.074 | 0.151 | 0.077 | 0.700 | 0.042 |
| 4 | 0.063 | 0.097 | 0.034 | 0.425 | 0.091 | 0.129 | 0.038 | 0.345 | 0.006 |
| 5 | 0.063 | 0.110 | 0.047 | 0.585 | 0.091 | 0.142 | 0.051 | 0.463 | 0.019 |
| 6 | 0.059 | 0.141 | 0.082 | 1.025 | 0.087 | 0.159 | 0.072 | 0.654 | 0.054 |
| 7 | 0.024 | 0.073 | 0.049 | 0.612 | 0.011 | 0.139 | 0.128 | 1.163 | 0.049 |

Розглянемо тепер випадок, коли функції належності $\mu_G(y)$ мають асиметричний характер, причому при $y > y_0$ крива $\mu_G(y)$ спадає більш повільно, ніж вона зростає при $y < y_0$ (рис.12.12). Це відповідає ситуації, коли менеджмент не відчуває суворих обмежень в ресурсах, допускає значні перевищення $y$ над $y_0$, а основні вимоги ставить до зменшення ризику втрати інформації.

*Варіант 6.*

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

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

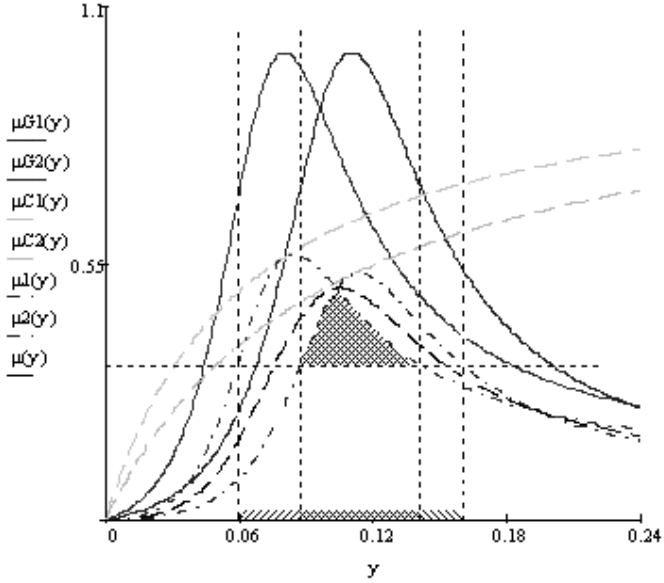



Рис. 12.12 . Функції належності при асиметричному характері $\mu_G(y)$

Наступний варіант відноситься до ситуації, коли менеджмент має протилежні пріоритети: витрачати значні кошти не має сенсу, така ситуація спостерігається, наприклад, коли 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. 12.13).

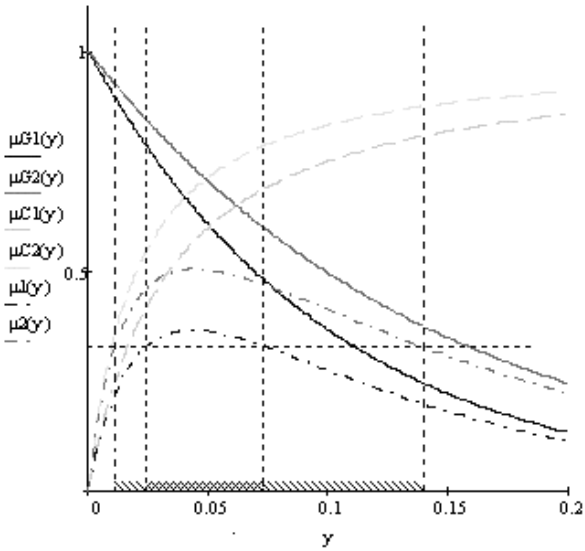



Рис. 12.13 . Функції належності при асиметричному характері $\mu_G(y)$

*Варіант 7.*

$$\mu_G{}^{(1)}(y)=e^{-10y}\quad \mu_C{}^{(1)}(y)=\frac{30y}{1+30y}$$

$$\mu_G{}^{(2)}(y)=e^{-7y}\quad \mu_C{}^{(2)}(y)=\frac{50y}{1+50y}$$

Відносно показаних на рисунках інтервалів перекриття зазначимо наступне. Якщо інтервали $\Delta$ 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 слугувати 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 $\Delta y^{(1)} = 0.046$, $\Delta y^{(2)} = 0.050$ у варіанті 1 до $\Delta y^{(1)} = 0.023$, $\Delta y^{(2)} = 0.025$ у варіанті 2;

2) перехід від дрібно-лінійної до дрібно-нелінійної функції належності $\mu_C(y)$ (варіант 5) приводить до незначного розширення інтервалів $\Delta y^{(k)}$, а саме - до $\Delta y^{(1)} = 0.047$, $\Delta y^{(2)} = 0.051$;

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:

$$\Delta y^{(1)} = 0.141 - 0.059 = 0.082$$
$$\Delta y^{(2)} = 0.159 - 0.087 = 0.072 \quad \text{у варіанті 6,}$$
$$\Delta y^{(1)} = 0.073 - 0.024 = 0.049$$
$$\Delta y^{(2)} = 0.139 - 0.011 = 0.128 \quad \text{у варіанті 7.}$$

Форма функцій належності характеризує рівень строгості менеджменту до поставлених умов. Зокрема, асиметричний вид функцій $\mu_G(y)$ відображає той факт, що більш високі вимоги ставляться до кількості захищеної інформації ніж до витрат на її захист. Експоненціальний характер $\mu_G(y)$ характерний для систем з дуже високою або дуже низькою вразливістю: в обох випадках витрачати невелику кількість ресурсів на захист недоцільно. Функції належності $\mu_C(y)$ визначаються вразливістю об'єктів.

Вище ми проаналізували походження функцій належності $\mu_C(y)$, зв'язавши їх з функціями $f(x, y)$, які використовуємо при «чіткому» підході. Звернемось тепер до функцій належності $\mu_G(y)$ і спробуємо з'ясувати, з яких міркувань менеджмент буде обирати

значення $y_0$ (в наших розрахунках вони залишились незмінними - $y_0^{(1)} = 0{,}08$, $y_0^{(2)} = 0{,}11$). На наш погляд, ці значення повинні обиратись з умови досягнення максимальної рентабельності інвестицій в захист інформації.

Рентабельність r визначається як відношення прибутку b до витрат. Прибуток в нашому випадку – це вартість захищеної інформації за відрахуванням витрат, які забезпечують цей захист.

Отже, $r(x,y) = \dfrac{b(x,y)}{y}$, де $b(x,y) = 1 - i(x,y) - y$.

Нагадаємо, що всі величини в наших розрахунках віднесені до вартості всієї інформації, яку ми покладаємо рівній 1, і при прийнятих допущеннях p=1, g=1 маємо $i(x,y) = f(x,y)$.

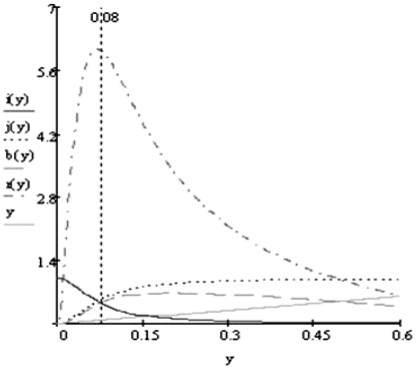
Припустимо, що вразливості об'єктів виражаються функціями

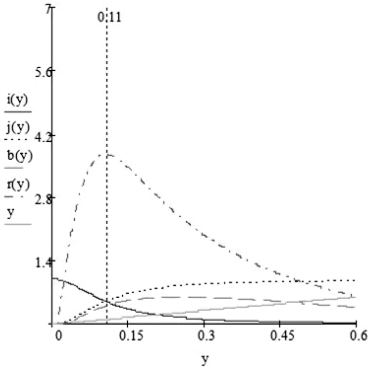
$$f^{(k)}(x,y) = \frac{\left(x/y\right)^2}{\left(x/y\right)^2 + c_{(k)}}.$$

Результати розрахунків, представлені на рис.11.14, засвідчують, що максимальна рентабельність інвестицій у захист інформації на двох об'єктах досягається при $y^{(1)} = 0{,}08$, $y^{(2)} = 0{,}11$ за умови використання функцій

$$f(x,y) = \frac{\left(x/y\right)^2}{\left(x/y\right)^2 + 200}, \quad f(x,y) = \frac{\left(x/y\right)^2}{\left(x/y\right)^2 + 90} \qquad (12.12)$$

Отже, можемо вважати «чіткий» підхід до системи, в якій вразливості двох об'єктів відображаються функціями (12.12), відповідає нечіткому підходу, в якому функції належності виражені у варіанті 5.




12.14,а 12.14,б

Рис. 12.14,а,б. Показники двох об’єктів в залежності від розміру інвестицій.

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 і розрахувати допустимі інтервали ресурсів захисту, які задовольняють поставленій меті і поставленому обмеженню для кожного об’єкта при різних формулюваннях функцій належності (ці інтервали показані на рисунках ліво- і право направленою штриховкою). Інтервал допустимих значень *y* для всієї системи визначається областю, в якій результуючий показник $\mu(y) = \sqrt{\mu^{(1)}(y) \cdot \mu^{(2)}(y)}$ (жирна штрихова лінія) перевищує заданий рівень $\mu = 0{,}33$.

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.

### Контрольні питання

1. Основні положення нечіткої логіки і теорії нечітких множин.
2. Сутність методу аналізу ієрархій.
3. Принципи нечіткого багатокритеріального аналізу об'єктів захисту інформації.
4. Сутність методики Белмана-Заде в застосуванні до задач інформаційної безпеки.
5. Визначення інтервалів допустимих втрат на основі нечіткого підходу.

## XIII. ВИЗНАЧЕННЯ СТАНІВ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

В попередніх розділах ми розглядали поодинокі атаки на інформаційну систему і визначали результати кожного протистояння, а саме значення таких показників, як кількість вилученої інформації, прибуток від інвестицій в інформаційну безпеку, рентабельність інвестицій – в залежності від умов розподілу інформації по об'єктах, їх вразливості, кількості ресурсів нападу і захисту. В реальних умовах атаки здійснюються постійно, і нас цікавить стан інформаційної безпеки, тобто значення приведених показників в будь-який момент часу.

### 13.1 Методика розрахунку перехідних ймовірностей і станів

Оптимізація показників є, безумовно, важливою задачею інформаційного менеджменту. Проте рішення цієї задачі направлене на досягнення певного результату (наприклад, частки захищеної інформації) в окремому акті протистояння двох сторін і не враховує статистику нападів. Таким чином, ми приходимо до необхідності визначення станів інформаційної безпеки в динаміці з врахуванням зміни параметрів протистояння з часом.

Будемо розглядати послідовні спроби вилучення інформації як марковський випадковий процес з дискретними станами і дискретним часом, тобто дискретний марковський ланцюг. Методику визначення станів на основі використання марковських

ланцюгів проілюструємо на найпростішому прикладі. Розглянемо об'єкт, для якого $x/y = 0...3,\ q(x,y) = \frac{1}{3},\ f(x,y) = \frac{x/y}{x/y + 2}$.

Поклавши *g=1, y=1,* маємо

$$f(x) = \frac{x}{x+2},\ i(x) = \frac{1}{3} \cdot \frac{x}{x+2} \qquad (13.1)$$

Три спроби вилучення інформації приводять до чотирьох можливих станів системи, які визначимо наступними умовами:

$s_1$ – інформація збережена повністю;

$s_2$ – втрачено $0<i\le 0,05$ всієї інформації;

$s_3$ – частка втраченої інформації лежить в межах $0,05<i\le 0,15$;

$s_4$ –частка втраченої інформації задовольняє умові $i>0,15$.

Перехідні імовірності $P_{ij}(k)$ марковського ланцюга, які характеризують перехід систем з *i*-го в *j*-тий стан на *k*-му кроці визначаються видом залежності *i(x)*. Вважаючи *x* неперервною випадковою величиною, визначимо $P_{12}(1)$ як геометричну імовірність потрапляння точки на відрізок $\Delta x_{12}^{(1)}$, котрий відповідає значенням $i\le 0,05$ (1). Граничне значення $x_{12}^{(1)}$ інтервалу $\Delta x_{12}^{(1)}$ знаходимо з умови $i(x_{12}^{(1)}) = \frac{1}{3} \cdot \frac{x_{12}^{(1)}}{x_{12}^{(1)} + 2} = 0,05$, звідки $x_{12}^{(1)} = 0,36$, $\Delta x_{12}^{(1)} = 0,36$. Оскільки увесь інтервал можливих значень *x* за нашим припущенням становить $\Delta x=3$, одержуємо $P_{12}(1) = \frac{\Delta x_{12}^{(1)}}{\Delta x} = \frac{0,36}{3} = 0,12$. Для розрахунку перехідної імовірності $P_{13}(1)$ знаходимо праву межу інтервалу $x_{13}^{(1)}$ (ліва межа інтервалу $\Delta x_{13}^{(1)}$ співпадає з правою межею інтервалу $\Delta x_{12}^{(1)}$). Покладаємо $\frac{1}{3} \cdot \frac{x_{13}^{(1)}}{x_{13}^{(1)} + 2} = 0,15$ і одержуємо $x_{13}^{(1)} = 1,67$, звідки:

$$P_{13}(1)=\frac{\Delta x_{13}^{(1)}}{\Delta x}=\frac{x_{13}^{(1)}-x_{12}^{(1)}}{\Delta x}=\frac{1{,}67-0{,}36}{3}=0{,}44.$$

Розраховуємо $P_{14}(1)$:

$$P_{14}(1)=\frac{\Delta x_{14}^{(1)}}{\Delta x}=\frac{x_{14}^{(1)}-x_{13}^{(1)}}{\Delta x}=\frac{3-1{,}67}{3}=0{,}44.$$

Використовуючи формулу повної імовірності, знаходимо:

$$P_{11}(1)=1-\left(P_{12}(1)+P_{13}(1)+P_{14}(1)\right)=0.$$

Перехідні імовірності на першому кроці утворюють матрицю-рядок: $P_{ij}(1)=\left(0 \quad 0{,}12 \quad 0{,}44 \quad 0{,}44\right)$.

Одержані результати можна отримати і графічно, знаходячи межі інтервалів, котрі визначаються граничними значеннями станів. Інтервали $\Delta x_{ij}(1)$ для $j$=2,3,4, показані горизонтальними лініями на рис.13.1, одержані з допомогою кривої $i_1$.

Зазначимо, що процеси, які відбуваються при інформаційному протистоянні, є однонаправленими, оскільки дії скеровані на вилучення інформації і інформація не поповнюється. Тому переходи з вищих станів в нижчі неможливі, зокрема $P_{21}(2)$=0. З цієї причини матриця перехідних ймовірностей є трикутною.

При розрахунку ймовірностей $P_{ij}(2)$ на другому кроці слід враховувати, що частина інформації вже вилучена на першому кроці. Після першого кроку об'єкт може знаходитись в одному з трьох станів: другому, третьому або четвертому. Отже, на другому кроці необхідно визначити перехідні імовірності $P_{22}(2)$, $P_{23}(2)$, $P_{24}(2)$,$P_{33}(2)$, $P_{34}(2)$, $P_{44}(2)$.

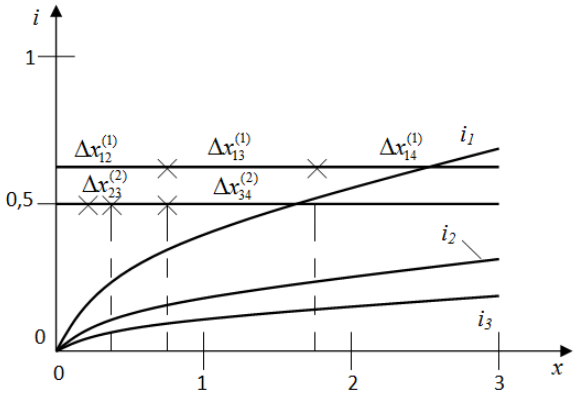


Рис. 13.1. Втрати інформації в залежності від ресурсів нападу

Вважатимемо, що на кожному кроці напад втрачає одну і ту ж кількість ресурсів. Тоді імовірність $P_{ij}(k)$ буде визначатись з одночасного виконання наступних умов:

1) Після $(k\text{-}1)$-го кроку об'єкт знаходився в $i$-му стані;

2) Після $k$-го кроку об'єкт має знаходитись в $j$-му стані;

3) Кількість інформації, втраченої під час всіх $k$спроб, має задовольняти умовам стану.

Наприклад, при розрахунку імовірності $P_{23}(2)$ ці умови приймають такий вигляд:

1) Після першого кроку об'єкт знаходився в другому стані, тобто знаходження $x$має лежати в інтервалі $\Delta x=0\ldots0{,}36$;

2) Після другого кроку об'єкт повинен знаходитись в третьому стані;

3) Кількість інформації, втраченої під час двох спроб задовольняє умові $0{,}05<i_1(x)+i_2(x)\leq 0{,}15$.

Використовуючи останню умову, знайдемо ліве граничне значення інтервалу $\Delta x_{23}^{(2)}$, котре співпадає з правим граничним значенням $x_{22}^{(2)}$ інтервалу $\Delta x_{22}^{(2)}$: $i_1(x_{22}^{(2)})+i_2(x_{22}^{(2)})=0{,}05$, звідки, спрощуючи запис, одержуємо:

$$i_1+(1-i_1)\cdot qf=qf+(1-qf)qf=2qf-q^2f^2=\frac{2}{3}f-\frac{1}{9}f^2=0{,}05$$

Розв'язуючи квадратне рівняння маємо:

$f$=0,235; $\dfrac{x_{22}^{(2)}}{x_{22}^{(2)}+2}=0{,}235$; $x_{22}^{(2)}=0{,}16$.

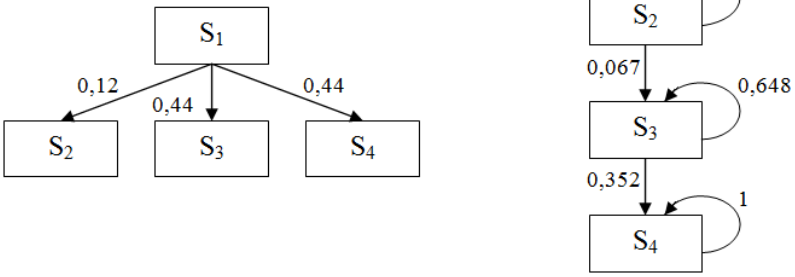
Праве граничне значення $x_{23}^{(2)}$ інтервалу $\Delta x_{23}^{(2)}$ визначається умовою: $\dfrac{2}{3}f-\dfrac{1}{9}f^2=0{,}15$, звідки знаходимо $x$=0,614.

Проте це значення виходить за межі дозволеного інтервалу $\Delta x$=0…0,36. Граничне значення $x_{23}^{(2)}$ буде обмежене шириною цього інтервалу, тобто $x_{23}^{(2)}=0{,}36$, $\Delta x_{23}^{(2)}=0{,}16...0{,}36$, звідки $P_{23}(2)=\dfrac{\Delta x_{23}(2)}{\Delta x}=\dfrac{0{,}36-0{,}16}{3}=0{,}067$. При перевищенні величини $x$=0,36 об'єкт після першої спроби знаходиться в третьому стані, і значення $x$ в інтервалі 0,36…1,67 характеризують переходи з третього в інші стани, тобто визначають імовірності $P_{33}(2)$, $P_{34}(2)$. Таким чином, імовірності $P_{22}^{(2)}$ і $P_{23}^{(2)}$ складають повну групу подій, тобто $P_{22}^{(2)}+P_{23}^{(2)}=1$, звідки $P_{22}^{(2)}=1-P_{23}^{(2)}=0{,}953$. Звертаючись до імовірності $P_{34}(2)$, врахуємо, що ліва межа інтервалу $\Delta x_{34}^{(2)}$ співпадає зі значенням $x$=0,614, а права обмежується граничним значенням інтервалу $\Delta x_{13}^{(1)}$ і становить $x$=1,67, оскільки стан $s_4$ не має верхньої межі. Отже, $P_{34}(2)=\dfrac{\Delta x_{34}^{(2)}}{\Delta x}=\dfrac{1{,}67-0{,}614}{3}=0{,}352$, а $P_{44}^{(2)}=1-P_{34}^{(2)}=0{,}648$.

Представимо матрицю $P_{ij}(2)$:

$$P_{ij}(2)=\begin{pmatrix}P_{11}(2) & P_{12}(2) & P_{13}(2) & P_{14}(2)\\ P_{21}(2) & P_{22}(2) & P_{23}(2) & P_{24}(2)\\ P_{31}(2) & P_{32}(2) & P_{33}(2) & P_{34}(2)\\ P_{41}(2) & P_{42}(2) & P_{43}(2) & P_{44}(2)\end{pmatrix}=\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 0{,}953 & 0{,}067 & 0\\ 0 & 0 & 0{,}648 & 0{,}352\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Графи станів для двох спроб зображені на рис. 13.2.




а) перша спроба а) друга спроба

Рис. 13.2. Графи станів для першої і другої спроб

Граф другої спроби має послідовний характер, тобто в ньому відсутній безпосередній перехід з другого стану в четвертий. Це є наслідком введеної нами умови: на кожному кроці втрачається однакова кількість ресурсів. Тому перехід з другого стану обмежений кількістю ресурсів $x_{12}^{(1)} = 0{,}36$, що недостатньо для переходу в четвертий стан.

Перехідні імовірності $P_{ij}(3)$ в результаті третьої спроби визначається через відповідні інтервали $\Delta x_{ij}^{(3)}$. Об'єкт перед третьою спробою може знаходитись в одному з трьох станів – другому, третьому або четвертому. Імовірності, які підлягають розрахунку – $P_{22}(3)$, $P_{23}(3)$, $P_{33}(3)$, $P_{34}(3)$, $P_{44}(3)$. Граничні значення $x_{ij}^{(3)}$ знаходимо з умови:

$$i_1(x_{ij}^{(3)}) + i_2(x_{ij}^{(3)}) + i_3(x_{ij}^{(3)}) = x_j,$$

де $x_j$ – граничне значення стану $s_j$. Візьмемо для прикладу $P_{23}(3)$. Тоді $i_1(x_{ij}^{(3)}) + i_2(x_{ij}^{(3)}) + i_3(x_{ij}^{(3)}) = 0{,}15$, або у скороченому вигляді:

$$i_1 + i_2 + i_3 = i_1 + (1 - i_1) \cdot qf + (1 - i_1 - i_2) \cdot qf = qf + (1 - qf) \cdot qf +$$

$$+ [1 - qf - (1 - qf) \cdot qf] = (qf)^3 - 3(qf)^2 + 3qf = 0{,}15$$

Після рішення кубічного рівняння з співвідношення $f = \dfrac{x_{23}^{(3)}}{x_{23}^{(3)} + 2}$ знаходимо $x_{23}^{(3)}$, а потім – і шукану величину $P_{23}(3)$.

Звичайно, зростання степеня алгебраїчного рівняння, яке необхідно розв'язати для знаходження значень $x_{ij}^{(k)}$, складає певні труднощі, проте на практиці кількість нападів, які супроводжуються вилученням інформації, не може бути значною.

Знайдемо тепер імовірності $p_j(k)$ станів об'єкта після кожної спроби. В початковий момент $p_1(0)=1$, $p_2(0)=p_3(0)=p_4(0)$ – об'єкт знаходиться в стані $s_1$. Імовірності станів після наступних спроб знаходимо за рекурентною формулою:

$$p_j(k) = \sum_{i=1}^{n} p_i(k-1)P_{ij}(k)$$

де $i, j = \overline{1,n}$ – номери станів після першої спроби визначаються значеннями $P_{ij}(1)$: $p_1(1)=0$; $p_2(1)=0{,}12$; $p_3(1)=0{,}44$; $p_4(1)=0{,}44$.

Після другої спроби маємо:

$$p_1(2) = p_1(1)P_{11}(2) + p_2(1)P_{21}(2) + p_3(1)P_{31}(2) + p_4(1)P_{41}(2) = 0$$

$$p_2(2) = p_1(1)P_{12}(2) + p_2(1)P_{22}(2) + p_3(1)P_{32}(2) + p_4(1)P_{42}(2) =$$

$$= 0{,}12 \cdot 0{,}953 = 0{,}114$$

$$p_3(2) = p_1(1)P_{13}(2) + p_2(1)P_{23}(2) + p_3(1)P_{33}(2) + p_4(1)P_{43}(2) =$$

$$= 0{,}12 \cdot 0{,}067 + 0{,}44 \cdot 0{,}648 = 0{,}293$$

$$p_4(2) = p_1(1)P_{14}(2) + p_2(1)P_{24}(2) + p_3(1)P_{34}(2) + p_4(1)P_{44}(2) =$$

$$= 0{,}44 \cdot 0{,}352 + 0{,}44 \cdot 1 = 0{,}595$$

## 13.2 Дискретні марковські ланцюги

### *13.2.1 Основні означення*

Будемо розглядати послідовні атаки як випадковий процес. Звернемось до особливо важливого і добре вивченого класу – марковських процесів[6][18], характерною особливістю котрих є відсутність післядії: імовірність будь-якого стану системи в

[6]А.А. Марков (1856 – 1922) –російський математик.

майбутньому (при $t>t_0$) залежить тільки від її стану в теперішньому ($t=t_0$) і не залежить від стану в минулому ($t<t_0$). Процеси з дискретними станами і дискретними переходами із стану в стан, називаються ланцюгами. Ланцюги, в яких майбутнє залежить від минулого лише через теперішнє, називаються марковськими

Перехід системи із стану $S_i$ в стан$S_j$ в момент часу$t$ визначається умовною імовірністю $P_{ij}(t)$. Матриця $P$, елементами котрої являються імовірності переходу системи за один крок, називається матрицею переходів:

$$P = \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1j} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2j} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P_{i1} & P_{i2} & \dots & P_{ij} & \dots & P_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nj} & \dots & P_{nn} \end{pmatrix}$$

Властивості матриці переходів:

1) Величини $P_{ij}$ визначають імовірність переходу з$i$-го в $j$-тий стан. Якщо індекси співпадають ($i=j$), то це імовірність того, що система залишається в і-му стані.

2) Оскільки переходи можливі тільки у майбутнє (в минуле неможливі), то при $i>j$ $P_{ij}=0$, і матриця стає трикутною.

3) Елементи кожного рядка визначають імовірності переходу зі-го стану в усі інші і створюють повну систему подій, тому сума елементів кожного рядка дорівнює 1:

$$\sum_{j=1}^{n} P_{ij} = 1\,,\ i = \overline{1,n},\quad P_{ii} = 1 - \sum_{\substack{j=1 \\ (\,j\neq i\,)}}^{n} P_{ij}$$

Разом з ймовірностями $P_{ij}$, які характеризують перехід системи з і-го в j-тий стан, введемо імовірності станів $p_j(k)$, які визначають імовірність того, що система через k кроків буде знаходитися в j-му стані. Ці величини, очевидно, задовольняють умові:

$$\sum_{j=1}^{n} p_j(k) = 1$$

В системах захисту інформації $p_j(k)$ визначають імовірність знаходження системи в стані, який характеризується певною кількістю збереженої інформації після декількох спроб її вилучення.

Імовірність стану $p_j(k)$ після k-го кроку, тобто k-ої спроби вилучення інформації визначається ймовірностями всіх інших станів на попередньому, k-1му кроці і ймовірностями переходів з усіх інших станів в j-тий за допомогою рекурентної формули:

$$p_j(k) = \sum_{i=1}^{n} p_i(k-1) \cdot P_{ij}, \qquad j = \overline{1,n}$$

Наприклад, імовірність знаходження системи в 3-му стані після 2-ої спроби при $n=4$визначається виразом:

$$p_3(2) = \sum_{i=1}^{4} p_i(1) \cdot P_{i3} = p_1(1) \cdot P_{13} + p_2(1) \cdot P_{23} + p_3(1) \cdot P_{33} + p_4(1) \cdot P_{43}$$

Для розрахунку ймовірностей станів $p_j(1)$після 1-го кроку необхідно задати початкові умови, тобто імовірності $p_j(0)$станів в початковій фазі.

Приведемо деякі визначення.

Якщо перехідні імовірності не залежать від номера кроку, то марковський ланцюг називається однорідним. Процес називається ергодичним, якщо система із будь-якого стану може перейти за скінченне число кроків в будь-який інший стан. Графом станів системи називають множину геометричних фігур (квадратів, кіл, прямокутників і т. д.), які умовно зображають стани, а стрілки між цими фігурами відображають безпосередні переходи зі стану в стан (рис. 13.3). Граф називається розміченим, якщо поряд зі стрілками приведені значення перехідних ймовірностей.

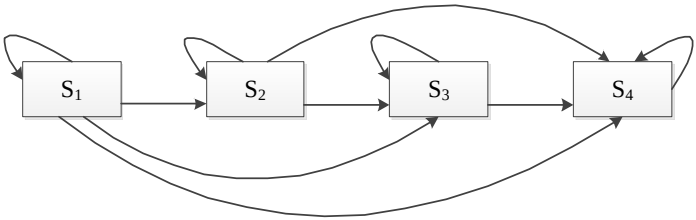



Рис.13.3. Граф станів.

### *13.2.2 Методика моделювання операцій при заданих значеннях перехідних ймовірностей.*

Розглянемо приклади.

*Приклад 1*. Суперник робить три спроби вилучення інформації. Можливі стани об'єкту інформаційної безпеки:

$S_1$ – $i = 0$

$S_2$ – $0 < i \le 0.33$

$S_3$ – $0.33 < i \le 0.5$

$S_4$ – $0.5 < i$

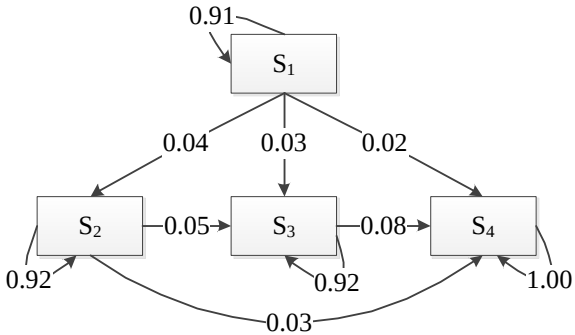



Рис. 13.4. Розмічений граф.

Імовірності $P_{ij}$ переходу об'єкта з одного стану в інший зазначені на графі (рис. 13.4). В початковий момент об'єкт знаходиться в стані $S_1$. Визначимо імовірності станів об'єкту після трьох спроб. Перехідні імовірності не залежить від номера кроку,

але переходи в попередні стани неможливі. Це однорідний неергодичний дискретний марковський ланцюг.

З графу станів знаходимо імовірності переходів після кожної спроби.

Імовірності переходів зі стану $S_1$ в інші стани після першої спроби зазначені на графі:

$P_{12} = 0,04;\ P_{13} = 0,03;\ P_{14} = 0,02.$

Використовуємо формулу повної імовірності і одержуємо значення $P_{11}$:

$P_{11} + P_{12} + P_{13} + P_{14} = 1 \qquad P_{11} = 1 - (P_{12} + P_{13} + P_{14}) = 0,91$

Імовірності переходів зі стану $\boldsymbol{S_2}$ в інші стани після другої спроби:

$P_{21} = 0$ (перехід зі стану $S_2$в стан $S_1$ неможливий)

$P_{23} = 0,05;\ P_{24} = 0,03;\ P_{22} = 1 - (P_{21} + P_{23} + P_{24}) = 0,92$

Аналогічно маємо для переходів зі стану $S_3$ і стану $S_4$:

$P_{31} = 0;\ P_{32} = 0;\ P_{34} = 0,08;\ P_{33} = 0,92$

$P_{41} = 0;\ P_{42} = 0;\ P_{43} = 0;\ P_{44} = 1$

Матриця перехідних ймовірностей має вигляд:

$$P = \begin{pmatrix} 0.91 & 0.04 & 0.03 & 0.02 \\ 0 & 0.92 & 0.05 & 0.03 \\ 0 & 0 & 0.92 & 0.08 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Згідно з умовою в початковий момент об'єкт знаходиться в стані $S_1$. Отже, $p_1(0) = 1;\ p_2(0) = p_3(0) = p_4(0) = 0.$

Імовірності станів об'єкту після першої спроби даються першим рядком матриці переходу:

$p_1(1) = 0,91;\ p_2(1) = 0,04;\ p_3(1) = 0,03;\ p_4(1) = 0,02$

Імовірності станів після другої спроби знаходимо за рекурентною формулою:

$p_j(k) = \sum_{i=1}^{n} p_i(k-1) \cdot P_{ij}$, $j = \overline{1,n}$ – номер стану; k – номер спроби.

Задаємо $k = 2$:

$$p_1(2) = p_1(1) \cdot P_{11} + p_2(1) \cdot P_{21} + p_3(1) \cdot P_{31} + p_4(1) \cdot P_{41} =$$
$$= 0.91 \cdot 0.91 + 0.04 \cdot 0 + 0.03 \cdot 0 + 0.02 \cdot 0 = 0.828$$
$$p_2(2) = p_1(1) \cdot P_{12} + p_2(1) \cdot P_{22} + p_3(1) \cdot P_{32} + p_4(1) \cdot P_{42} =$$
$$= 0.91 \cdot 0.04 + 0.04 \cdot 0.92 + 0.03 \cdot 0 + 0.02 \cdot 0 = 0.073$$
$$p_3(2) = p_1(1) \cdot P_{13} + p_2(1) \cdot P_{23} + p_3(1) \cdot P_{33} + p_4(1) \cdot P_{43} =$$
$$= 0.91 \cdot 0.03 + 0.04 \cdot 0.05 + 0.03 \cdot 0.92 + 0.02 \cdot 0 = 0.057$$
$$p_4(2) = p_1(1) \cdot P_{14} + p_2(1) \cdot P_{24} + p_3(1) \cdot P_{34} + p_4(1) \cdot P_{44} =$$
$$= 0.91 \cdot 0.02 + 0.04 \cdot 0.03 + 0.03 \cdot 0.08 + 0.02 \cdot 1 = 0.042$$

Задаючи $k = 3$, знаходимо:

$$p_1(3) = p_1(2) \cdot P_{11} + p_2(2) \cdot P_{21} + p_3(2) \cdot P_{31} + p_4(2) \cdot P_{41} =$$
$$= 0.828 \cdot 0.91 = 0.753$$
$$p_2(3) = p_1(2) \cdot P_{12} + p_2(2) \cdot P_{22} + p_3(2) \cdot P_{32} + p_4(2) \cdot P_{42} =$$
$$= 0.828 \cdot 0.04 + 0.073 \cdot 0.92 = 0.100$$
$$p_3(3) = p_1(2) \cdot P_{13} + p_2(2) \cdot P_{23} + p_3(2) \cdot P_{33} + p_4(2) \cdot P_{43} =$$
$$= 0.828 \cdot 0.03 + 0.073 \cdot 0.05 + 0.057 \cdot 0.92 = 0.081$$
$$p_4(3) = p_1(2) \cdot P_{14} + p_2(2) \cdot P_{24} + p_3(2) \cdot P_{34} + p_4(2) \cdot P_{44} =$$
$$= 0.828 \cdot 0.02 + 0.073 \cdot 0.03 + 0.057 \cdot 0.08 + 0.042 \cdot 1 = 0.066$$

Отже, маємо такі імовірності всіх можливих станів після трьох спроб вилучення інформації:

1) Інформація збережена повністю $i = 0$ $p_1(3) = 0{,}753$
2) Вилучена кількість інформації $0 < i \le 0.33$ $p_2(3) = 0{,}100$
3) Вилучена кількість інформації $0.33 < i \le 0.5$ $p_3(3) = 0{,}081$
4) Вилучена кількість інформації $0.5 < i$ $p_4(3) = 0{,}066$

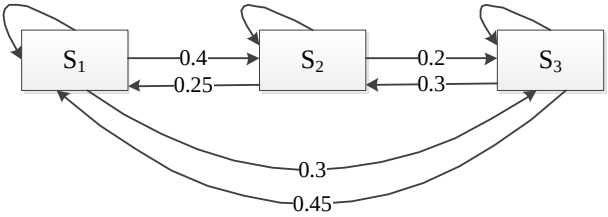
Оцінка рівня інформаційної безпеки на основі одержаних результатів залишається за менеджментом системи.

Марківський ланцюг називається неоднорідним, якщо хоча б одна з перехідних ймовірностей залежить від номера кроку.

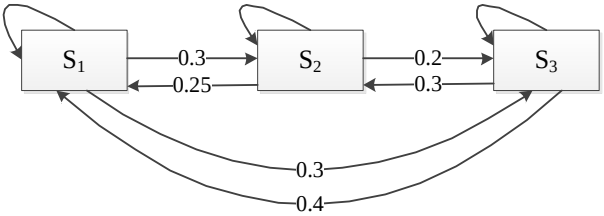
Матриця переходів і розмічений граф також будуть змінюватись від кроку до кроку:

$$P_{ij}(k) = \begin{pmatrix} P_{11}(k) & \dots & P_{1n}(k) \\ \dots & \dots & \dots \\ P_{n1}(k) & \dots & P_{nn}(k) \end{pmatrix}$$

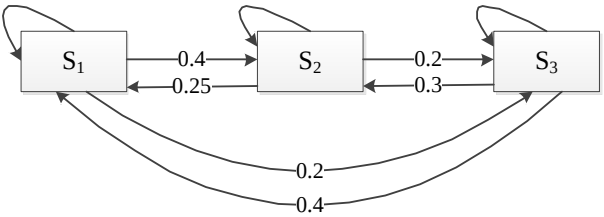
*Приклад 2*. Розмічений граф на першому кроці має такий вигляд:




На другому кроці:




На третьому кроці:




Імовірності $P_{ii}$ знаходимо з формули повної імовірності, а потім за приведеною методикою - імовірності станів з урахуванням зміни перехідних ймовірностей на кожному кроці.

### *12.2.3 Визначення перехідних ймовірностей*

Імовірність переходу системи з стану в стан залежить від частки інформації, котра може бути вилучена на кожному кроці. Ця величина, в свою чергу, визначається вразливістю об'єкта і відносною кількістю ресурсів нападу $\frac{x}{y}$. Для спрощення запису покладемо$y$=1. Почнемо з найпростішого випадку, коли імовірність виділення ресурсів нападу в певних межах (візьмемо$x = 0 \ldots 3$) вважається однаковою($q$=1), а динамічна вразливість задається дрібно-лінійною функцією

$$f(x) = \frac{x}{x+2}.$$

Частка інформації, вилученої на k-му кроці, становить

$$i_k(x) = g_k f_k(x) = g_k \frac{x}{x+2}.$$

На першому кроці , на наступних , оскільки враховується інформація, вилучена на попередніх кроках. Вважатимемо також, що значення$x$ на кожному кроці залишається незмінним.

Будемо розглядати послідовні спроби вилучення інформації як марківський випадковий процес з дискретними станами і дискретним часом. Обмежимось розглядом трьох спроб і чотирьох можливих станів, які визначимо наступними умовами:

$S_1$ – інформація збережена повністю;

$S_2$ – вилучено $i \le 0{,}33$ всієї інформації;

$S_3$– частка вилученої інформації лежить в межах$0{,}33 \le i < 0{,}5$;

$S_4$– частка вилученої інформації $i > 0{,}5$.

Перехідні імовірності $P_{ij}(k)$ марковського ланцюга поставленої задачі визначається залежностями$i_k(x)$ (рис.12.3). Почнемо з імовірності $P_{12}(1)$. Вважаючи$x$ неперервною випадковою величиною, визначимо $P_{12}(1)$ як геометричну імовірність потрапляння точки $x$ на відрізок$\Delta x_{12}$, котрий відповідає значенням $0 \le i(x) < 0{,}33$. Ліва границя цього відрізка $x_{12}^{(1)} = 0$, а праву границю

$x_{12}^{(2)}$ знайдемо з умови $-\dfrac{x_{12}^{(2)}}{x_{12}^{(2)}+2}=0{,}33$, звідки $x_{12}^{(2)}=1$ і $\Delta x_{12}$=0…1. Відношення цього інтервалу до загального інтервалу 0…3 і визначає $P_{12}(1)=\dfrac{\Delta x_{12}(1)}{\Delta x}=\dfrac{1}{3}=0{,}33$. Для інтервалу $\Delta x_{13}$ліва границя $x_{13}^{(1)}$ співпадає з границею $x_{12}^{(2)}=1$ попереднього інтервалу, а права визначається з рівності $\dfrac{x_{13}^{(2)}}{x_{13}^{(2)}+2}=0{,}5$, звідки $x_{13}^{(2)}=2$, $\Delta x_{13}$=1…2і$P_{13}(1)=\frac{\Delta x_{13}(1)}{\Delta x}=\frac{2-1}{3}=0{,}33$. Для $P_{14}(1)$маємо $x_{14}^{(1)}=2$, і $x_{14}^{(2)}=3$ $P_{14}(1)=\dfrac{\Delta x_{42}(1)}{\Delta x}=\dfrac{3-2}{3}=0{,}33$. Інтервали $\Delta x_{ij}$ для трьох послідовних спроб зображені на полі рис. 12.3. Зазначимо, що величини цих інтервалів можна визначити також графічно, користуючись кривою $i_1(x)$.

Почнемо з $P_{22}(2)$. Це імовірність того, що об'єкт, який знаходився в другому стані після першого кроку, залишився в цьому ж стані після другого кроку. Інакше кажучи, після здійснених двох кроків частка вилученої інформації повинна лежати в межах 0…0,33. Ліва границя інтервалу $\Delta x_{22}(2)$ дорівнює $x_{22}^{(2)}=0$, а праву знаходимо з умови

$$i_1(x_{22}^{(2)})+i_2(x_{22}^{(2)})=\frac{x_{22}^{(2)}}{x_{22}^{(2)}+2}+(1-\frac{x_{22}^{(2)}}{x_{22}^{(2)}+2})\frac{x_{22}^{(2)}}{x_{22}^{(2)}+2}=0{,}33$$

Враховуємо,що $\dfrac{x_{22}^{(2)}}{x_{22}^{(2)}+2}=f$ і знайдемо цю величину:

$$f+(1-f)f=2f-f^2=0{,}33$$

Розв'язуємо квадратне рівняння $f^2-2f+0{,}33=0$ і знаходимо $f_1$=1,82, $f_2$=0,18. Перший корінь не має фізичного смислу(він приводить до значення $x_{22}^{(2)}$<0), а другий дає $x_{22}^{(2)}=0{,}44$. Перевірка показує, що в цій точці$i_1$=0,18, $i_2$=0,15, $i=i_1+i_2$=0,33. Отже,

$x_{22}^{(2)} = 0{,}44$, $\Delta x_{22}^{(2)} = 0...0{,}44$, звідки $P_{22}(2) = \dfrac{\Delta x_{22}(2)}{\Delta x}$, де $x$=0…1 – весь інтервал можливих значень, який визначається умовою, що перехід відбувається з другого стану після першої спроби. Отже, $P_{22}(2) = \dfrac{0{,}44}{1} = 0{,}44$. Зауважимо, що цей розрахунок можна провести графічно, додаючи ординати залежностей $i_1(x)$ та $i_2(x)$ в різних точках і знаходячи точку $x_{22}^{(2)}$, в якій ця сума дорівнює 0,33.

При розрахунку ймовірностей $P_{23}(2)$ врахуємо, що права границя інтервалу $\Delta x_{22}(1)$ є одночасно лівою границею інтервалу $\Delta x_{23}(2)$, тобто $x_{23}^{(1)} = x_{22}^{(2)} = 0{,}44$. Праву границю $x_{23}^{(2)}$ інтервалу $\Delta x_{23}(2)$ знайдемо з рівняння $f + (1-f)f = 0{,}5$, звідки $f^2 - 2f + 0{,}5 = 0$, $f$=0,29, $x_{23}^{(2)} = 0{,}8$, $\Delta x_{23}^{(2)} = 0{,}8 - 0{,}44 = 0{,}36$,

$$P_{23}(2) = \frac{\Delta x_{23}(2)}{\Delta x} = \frac{0{,}36}{1} = 0{,}36.$$

Розрахунок ймовірностей $P_{24}(2)$ дещо відрізняється від попередніх розрахунків. Це пов’язано з тим, що в формулюванні $4^{\underline{го}}$ стану зазначена лише ліва границя$i$>0,5. Вона й визначатиме ліву границю інтервалу$\Delta x_{24}(2)$: $x_{24}^{(1)} = x_{23}^{(2)} = 0{,}8$. Права границя знаходиться з умови, що об’єкт після першої спроби знаходиться в $2^{\underline{му}}$ стані, а це значить,що $x$лежить в межах 0…1. Оскільки значення $x$від спроби до спроби не змінюються, то $x_{24}^{(2)} = 1$, $\Delta x_{24}^{(2)} = 1{,}0 - 0{,}8 = 0{,}2$, $P_{24}(2) = \dfrac{\Delta x_{24}(2)}{\Delta x} = \dfrac{0{,}2}{1} = 0{,}2$.

Після першого кроку об’єкт може знаходитись також в $3^{\underline{му}}$ і $4^{\underline{му}}$ станах. Розглядаючи переходи з $3^{\underline{го}}$ стану, необхідно розрахувати імовірності$P_{34}(2)$ і$P_{33}(2)$ . Інтервал можливих значень $x$ при переході з $3^{\underline{го}}$ стану лежить в межах$x$=1..2. Знайдемо$i=i_1(x)+i_2(x)$ на лівій границі інтервалу, тобто при $x$=1:

$$g_1 f(x) + g_2 f(x) = \frac{x}{x+2} + (1 - \frac{x}{x+2})\frac{x}{x+2} = \frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3} = 0{,}55$$

Ця точка відноситься до $4^{го}$ стану. Отже, переходи з усіх точок інтервалу $\Delta x$=1..2, тобто з $3^{го}$ стану можливі тільки в $4^{ий}$ стан: $P_{34}(2)=1$, $P_{33}(2)=0$. З цих же причин $P_{44}(2)=1$.

Побудуємо матриці перехідних ймовірностей і відповідні графи. На першому кроці переходи можливі з $1^{го}$ стану, тому елементи $Pij(1)$ формують матрицю-рядок

$$P_{ij}(1)=\begin{pmatrix}P_{11}(1) & P_{12}(1) & P_{13}(1) & P_{14}(1)\end{pmatrix}=\begin{pmatrix}0 & 0{,}33 & 0{,}33 & 0{,}33\end{pmatrix}.$$

Розмітимо граф цих переходів (рис. 12.4)

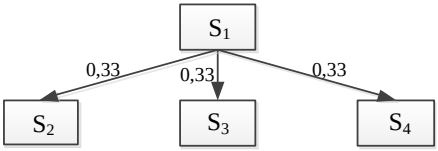



Рис. 13.5 Граф переходів на першому кроці.

На другому кроці об’єкт може знаходитись в $2^{му}$, $3^{му}$, $4^{му}$ стані. Отже, переходи можливі тільки з цих станів. Враховуючи одержані значення перехідних ймовірностей, будуємо матрицю

$$P_{ij}(2)=\begin{pmatrix}P_{21}(2) & P_{22}(2) & P_{23}(2) & P_{24}(2)\\ P_{31}(2) & P_{32}(2) & P_{33}(2) & P_{34}(2)\\ P_{41}(2) & P_{42}(2) & P_{43}(2) & P_{44}(2)\end{pmatrix}=\begin{pmatrix}0 & 0{,}44 & 0{,}36 & 0{,}20\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 1\end{pmatrix}$$

і відповідний граф (рис. 12.5)

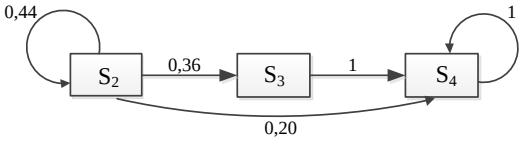



Рис. 13.6 Граф переходів на другому кроці.

## 13.3 Неперервні марківські ланцюги

### *13.3.1Основні означення*

Ми розглянули спроби вилучення інформації як кроки марківського ланцюга з дискретними станами і дискретним часом. В реальних ситуаціях ці спроби можуть здійснюватись не в певні дискретні, а й у будь-які моменти, тобто являють собою марковські випадкові процеси з дискретними станами і неперервним часом, або неперервні марківські ланцюги [4]. В цьому випадку стан інформаційної безпеки оцінюється не кількістю кроків, а при заданій інтенсивності потоку подій – часовою залежністю ймовірностей переходу системи зі стану в стан.

Позначимо $p_i(t)$ – імовірність того, що в момент $t$ система знаходиться в стані $S_i (i=1,2...n)$. Оскільки в будь-який момент $t$ система буде знаходитись в одному зі станів $S_1, S_2...S_n$, то події $S_i$, які відповідають цим станом, несумісні і утворюють повну групу. Тому має місце умова:

$$\sum_{i=1}^{n} p_i(t) = 1$$

В процесах з неперервним часом перехідні імовірності $P_{ij}$ уже не можуть бути характеристиками процесу: імовірність переходу системи з одного стану в інший точно в момент $t$ дорівнює нулю (так само, як імовірність будь-якого окремого значення неперервної випадкової величини). Це пов'язано з тим, що на вісі часу момент $t$ позначається точкою, яка не має розміру. Імовірність переходу ми можемо визначити лише для певного проміжку часу $\Delta t$. Тому вводять поняття щільності імовірності переходу зі стану $S_i$ в стан $S_j$.

$$\lambda_{ij}(t) = \lim_{\Delta t \to 0} \frac{P_{ij}(t;\Delta t)}{\Delta t}$$

При малому $\Delta t$ імовірність переходу $P_{ij}(t;\Delta t)$ з точністю до нескінченно малих вищих порядків дорівнює:

$$P_{ij}(t;\Delta t) \cong \lambda_{ij}(t) \cdot \Delta t$$

Якщо при будь-яких $i \neq j$ щільності ймовірностей переходів не залежать від часу, то процес називається однорідним і замість $\lambda_{ij}(t)$ пишемо просто $\boldsymbol{\lambda_{ij}}$. Величини $\boldsymbol{\lambda_{ij}}$утворюють матрицю:

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2n} \\ \dots & \dots & \dots & \dots \\ \lambda_{n1} & \lambda_{n2} & \dots & \lambda_{nn} \end{pmatrix}$$

Імовірності станів $p_i(t)$ знаходяться з системи диференціальних рівнянь Колмогорова[7]:

$$\frac{dp_i(t)}{dt} = -\left(\sum_{j=1}^{n} \lambda_{ij}\right) \cdot p_i(t) + \sum_{j=1}^{n} \lambda_{ij} \cdot p_i(t) \qquad i = 1,2...n; t \geq 0$$

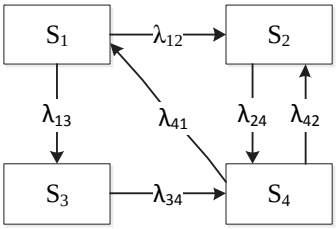
Структура рівнянь Колмогорова дуже проста. В лівій частині рівняння стоїть похідна імовірності стану, а права містить стільки членів, скільки стрілок на графі пов'язано з даним станом. Якщо стрілка направлена зі стану, то відповідний член має знак «мінус», якщо в стан - знак «плюс». Кожний член являє собою добуток щільності імовірності переходу, який відповідає даній стрілці, на імовірність стану, з якого виходить стрілка.

### *13.3.2 Методика моделювання операцій за даними розміченого графа*

Принципи моделювання операцій розглянемо на прикладах.

*Приклад 3.* Побудувати систему диференціальних рівнянь Колмогорова, яка відповідає розміченому графу:

[7] А. М. Колмогоров (1903 – 1987) радянський математик




$$\frac{dp_1(t)}{dt} = -(\lambda_{12} + \lambda_{13}) \cdot p_1(t) + \lambda_{41} \cdot p_4(t)$$

$$\frac{dp_2(t)}{dt} = -\lambda_{24} \cdot p_2(t) + \lambda_{12} \cdot p_1(t) + \lambda_{42} \cdot p_4(t)$$

$$\frac{dp_3(t)}{dt} = -\lambda_{34} \cdot p_3(t) + \lambda_{13} \cdot p_1(t)$$

$$\frac{dp_4(t)}{dt} = -(\lambda_{41} + \lambda_{42}) \cdot p_4(t) + \lambda_{24} \cdot p_2(t) + \lambda_{34} \cdot p_3(t)$$

*Приклад 4.*Розглянемо систему захисту приміщення, яка може знаходитись в одному з таких станів:

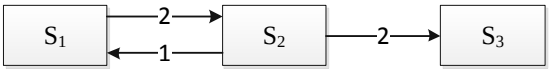
$S_1$ -система справна, але не включена;

$S_2$- система справна і включена;

$S_3$- система несправна

Будемо вважати, що система може вийти з ладу тільки під час експлуатації, тобто безпосередній перехід зі стану $S_1$в стан $S_3$неможливий. Вважатимемо також, що система може переходити зі стану в стан в будь-який випадковий момент часу, причому щільності ймовірностей переходів $\boldsymbol{\lambda_{ij}}$не залежать від часу. Отже, ми маємо однорідний марковський випадковий процес з неперервним часом.

Припустимо, що розмічений граф станів системи має такий вигляд:




Зазначимо, що щільності $\lambda_{ij}$, на відміну від перехідних ймовірностей дискретних ланцюгів, можуть приймати значення $\lambda > 1$.Як видно з графа, перехід системи зі стану $S_3$ в стан $S_2$(тобто ремонт системи) за проміжок часу, що розглядається, не

передбачено. Такий стан називають пасткою. Необхідно знайти імовірності станів в моменті $t = 1$, якщо в моменті $t = 0$ система знаходилась в стані $S_1$.

Матриця щільностей ймовірностей переходів, складена по приведеному графу, має вигляд:

$$\Lambda = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Знайдемо спочатку імовірності станів $p_1(t)$, $p_2(t)$, $p_3(t)$ в будь-який момент часу $t$. Складемо систему диференціальних рівнянь Колмогорова:

$$\begin{cases} \dfrac{dp_1(t)}{dt} = -2 \cdot p_1(t) + p_2(t) \\ \dfrac{dp_2(t)}{dt} = -3 \cdot p_2(t) + 2 \cdot p_1(t) \\ \dfrac{dp_3(t)}{dt} = 2 \cdot p_2(t) \end{cases} \quad (1)$$

Початкові умови: $p_1(0) = 1$; $p_2(0) = p_3(0) = 0$.

Перші два рівняння системи не містять невідому функцію $p_3(t)$, і тому їх можна розглядати як систему двох рівнянь з двома невідомими $p_1(t)$ і $p_2(t)$. Розв'яжемо цю систему:

$$\begin{cases} \dfrac{dp_1(t)}{dt} + 2 \cdot p_1(t) - p_2(t) = 0 \\ \dfrac{dp_2(t)}{dt} + 3 \cdot p_2(t) - 2 \cdot p_1(t) = 0 \end{cases} \quad (2)$$

Маємо систему лінійних диференціальних рівнянь першого порядку з постійними коефіцієнтами. Розв'язок шукаємо у вигляді показникових функцій:

$$p_1(t) = \gamma_1 \cdot e^{\mu \cdot t}$$

$$p_2(t) = \gamma_2 \cdot e^{\mu \cdot t}$$

де $\gamma_1$, $\gamma_2$, $\mu$ – шукані константи. Підставимо ці функції в систему і скоротимо на $e^{\mu t} > 0$. Одержимо:

$$\begin{cases}(2+\mu)\cdot\gamma_1 - \gamma_2 = 0\\ -2\cdot\gamma_1 + (3+\mu)\cdot\gamma_2 = 0\end{cases} \quad (3)$$

Ненульовий розв'язок системи однорідних алгебраїчних рівнянь існує тільки тоді, коли її визначник дорівнює нулю:

$$\begin{vmatrix}2+\mu & -1\\ -2 & 3+\mu\end{vmatrix} = 0$$

Це характеристичне рівняння системи диференціальних рівнянь, з якого знаходимо власні числа $\mu$.

Розкриваємо визначник:

$$(2+\mu)\cdot(3+\mu) - 2 = 0;$$

$$6 + 3\cdot\mu + 2\cdot\mu + \mu^2 - 2 = 0;$$

$$\mu^2 + 5\cdot\mu + 4 = 0$$

$$\mu_{1,2} = \frac{-5 \pm \sqrt{25-16}}{2} = \frac{-5\pm 3}{2}$$

$$\mu_1 = -4$$

$$\mu_2 = -1$$

Система має розв'язок лише при значеннях $\mu = \mu_1$і $\mu = \mu_2$. Підставимо в систему спочатку $\mu = \mu_1$і знайдемо значення $\gamma_1^{(1)}$і $\gamma_2^{(1)}$, вірніше, співвідношення між ними:

$$-2\cdot\gamma_1^{(1)} - \gamma_2^{(1)} = 0 \Rightarrow \gamma_2^{(1)} = -2\cdot\gamma_1^{(1)}$$

Тут $\gamma_1$ – вільне невідоме, тобто воно може приймати будь-яке значення. Покладемо $\gamma_1^{(1)} = 1$. Тоді $\gamma_2^{(1)} = -2$ і перший частинний розв'язок системи диференціальних рівнянь має вигляд:

$$p_1^{(1)}(t) = \gamma_1^{(1)}\cdot e^{\mu_1\cdot t} = e^{-4\cdot t}$$

$$p_2^{(1)}(t) = \gamma_2^{(2)}\cdot e^{\mu_1\cdot t} = -2\cdot e^{-4\cdot t}$$

Тепер підставимо в систему $\mu = \mu_2$і знайдемо другий частинний розв'язок:

$$\gamma_1^{(2)} - \gamma_2^{(2)} = 0 \Rightarrow \gamma_1^{(2)} = \gamma_2^{(2)}$$

Покладемо $\gamma_1^{(2)} = 1$. Тоді $\gamma_2^{(2)} = 1$,і маємо другий частинний розв'язок:

$$p_1^{(2)}(t) = \gamma_1^{(2)} \cdot e^{\mu_2 \cdot t} = e^{-t}$$

$$p_2^{(2)}(t) = \gamma_2^{(2)} \cdot e^{\mu_2 \cdot t} = e^{-t}$$

З частинних розв'язків складаємо загальний розв'язок системи:

$$\begin{cases} p_1(t) = C_1 \cdot p_1^{(1)}(t) + C_2 \cdot p_1^{(2)}(t) = C_1 \cdot e^{-4 \cdot t} + C_2 \cdot e^{-t} \\ p_2(t) = C_1 \cdot p_2^{(1)}(t) + C_2 \cdot p_2^{(2)}(t) = -2 \cdot C_1 \cdot e^{-4 \cdot t} + C_2 \cdot e^{-t} \end{cases} \quad (4)$$

де $C_1$ і $C_2$ – довільні константи, які знаходимо з початкових умов:

$$\begin{cases} p_1(0) = C_1 + C_2 = 1 \\ p_2(0) = -2 \cdot C_1 + C_2 = 0 \end{cases}$$

$$\begin{cases} C_2 = 1 - C_1 \\ -2 \cdot C_1 + C_2 = 0 \end{cases} \Rightarrow C_1 = \frac{1}{3} \quad C_2 = \frac{2}{3}$$

Підставляючи ці значення в систему (4), одержуємо шуканий частинний розв'язок, який задовольняє початковим умовам:

$$\begin{cases} p_1(t) = \frac{1}{3} \cdot e^{-4 \cdot t} + \frac{2}{3} \cdot e^{-t} \\ p_2(t) = -\frac{2}{3} \cdot e^{-4 \cdot t} + \frac{2}{3} \cdot e^{-t} \end{cases} \quad (5)$$

Функцію $p_3(t)$можна знайти з третього рівняння системи (1), але простіше скористатись нормовочною умовою для довільного ***t***:

$$\sum_{i=1}^{n} p_i(t) = 1$$

$$p_3(t) = 1 - p_1(t) - p_2(t) = 1 - \frac{1}{3}\cdot e^{-4t} - \frac{2}{3}\cdot e^{-t} + \frac{2}{3}\cdot e^{-4t} - \frac{2}{3}\cdot e^{-t} =$$

$$= \frac{1}{3}\cdot e^{-4t} - \frac{4}{3}\cdot e^{-t} + 1 \qquad (6)$$

Використовуючи (5), (6), визначимо шукані імовірності станів в момент $t = 1$:

$$\begin{cases} p_1(1) = \frac{1}{3}\cdot e^{-4} + \frac{2}{3}\cdot e^{-1} = 0.251 \\ p_2(1) = -\frac{2}{3}\cdot e^{-4} + \frac{2}{3}\cdot e^{-1} = 0.233 \\ p_3(1) = \frac{1}{3}\cdot e^{-4} - \frac{4}{3}\cdot e^{-1} + 1 = 0.516 \end{cases}$$

Висновок: якість системи незадовільна: $p_3 > p_1 > p_2$.

### *13.3.3 Визначення станів в багаторубіжній системі*

Розглянемо інформаційну систему, яка складається з трьох однакових об'єктів, захищених чотирма перешкодами, розташованими за послідовно-паралельною схемою (рис. 13.7).

Одна з перешкод є загальною (це може бути периметр території, що охороняється), інші – індивідуальні (окремі приміщення). Кожен з об'єктів містить об'єм інформації $g$. Через $X$ і $Y$ позначено загальна кількість ресурсів нападу і, відповідно, захисту.

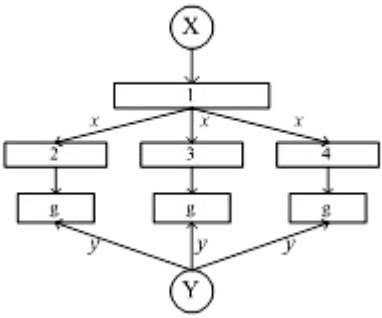



Рис. 13.7 Розташування перешкод в системі захисту інформації

На подолання кожної з перешкод напад виділяє кількість ресурсів $x$ ($X=4x$), на захист кожного з об'єктів – кількість ресурсів $y$ ($Y=3y$).

Вважатимемо, що напади здійснюються послідовно, утворюючи ординарний пуассонівський випадковий потік, який формує неперервний марковський ланцюг. Протистояння відбувається за такою схемою. Напад спрямовується спочатку на першу перешкоду, а після її подолання напади розподіляються рівномірно на подолання всіх інших перешкод. Стани інформаційної системи визначимо наступним чином:

$S_1$ – вся система непорушна;

$S_2$ – подолана тільки перша перешкода (інформація ще не вилучена);

$S_3$– подолана одна з індивідуальних перешкод, вилучена інформація кількістю $g$ з одного об'єкта;

$S_4$– подолані дві індивідуальні перешкоди, вилучено $2g$ інформації;

$S_5$– подолані всі перешкоди, вилучена вся інформація кількістю $3g$.

Позначимо: $\lambda$ – інтенсивність нападів, тобто їх кількість в одиницю часу, $p$ – імовірність того, що напад буде успішним і перешкода буде подолана. Тоді $\lambda p$ – інтенсивність успішних нападів, яка в нашій системі виражає щільність імовірності $\lambda_{ij}$ переходу систем з $i$-го в $j$-ий стан. За рахунок рівномірності зміни станів ця величина однакова для всіх переходів: $\lambda_{ij}= \lambda p$.

Граф сформульованої задачі завдяки ординарності пуассонівського потоку носить не розгалужений, а послідовний характер (рис. 13.8).

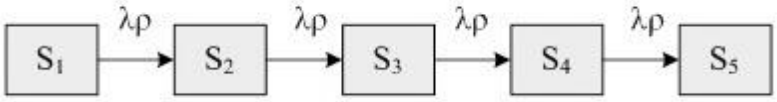



Рис. 13.8. Граф станів системи.

Приведеному графу відповідає система рівнянь Колмогорова:

$$\frac{dp_1}{dt} = -\lambda p p_1$$
$$\frac{dp_2}{dt} = -\lambda p p_2 + \lambda p p_1$$
$$\frac{dp_3}{dt} = -\lambda p p_3 + \lambda p p_2$$
$$\frac{dp_4}{dt} = -\lambda p p_4 + \lambda p p_3$$
$$\frac{dp_5}{dt} = \lambda p p_5.$$

*Приклад 5.*Початкові умови:

$$p_1(0) = 1, p_2(0) = p_3(0) = p_4(0) = p_5(0) = 0.$$

В цій системі імовірність $p$ досягнення результату при нападі – величина відома, а імовірності станів $p_i$ - невідомі, які необхідно визначити. Розв'язок почнемо з першого рівняння, яке є відокремленим, оскільки містить лише одну невідому – $p_1$. Застосуємо перетворення Лапласа:

$$p_1(t) \to \tilde{p}_1(\nu).$$

Змінну в перетворенні Лапласа позначимо через $\nu$, оскільки через $p$ ми позначаємо імовірність. Перше рівняння після перетворення має вигляд:

$$\tilde{p}_1(\nu) - p_1(0) = -\lambda p \tilde{p}_1(\nu) \Rightarrow \tilde{p}_1(\nu) = \frac{1}{\nu + \lambda p}, \quad p_1(t) = e^{-\lambda p t}.$$

Друге рівняння після перетворення має вигляд:

$$\tilde{p}_2(\nu) - p_2(0) = -\lambda p \tilde{p}_2(\nu) + \frac{\lambda p}{\nu + \lambda p} \Rightarrow \tilde{p}_2(\nu) = \frac{\lambda p}{(\nu + \lambda p)^2},$$

$$p_2(t) = \frac{\lambda p t}{1!} e^{-\lambda p t}.$$

Аналогічно маємо для наступних рівнянь:

$$\tilde{p}_3(\nu) - p_3(0) = -\lambda p \tilde{p}_3(\nu) + (\frac{\lambda p}{\nu + \lambda p})^2 \Rightarrow \tilde{p}_3(\nu) = (\frac{\lambda p}{\nu + \lambda p})^2,$$

$$p_3(t) = \frac{(\lambda pt)^2}{2!} e^{-\lambda pt}.$$

$$\tilde{p}_4(\nu) - p_4(0) = -\lambda p \tilde{p}_4(\nu) + (\frac{\lambda p}{\nu + \lambda p})^3 \Rightarrow \tilde{p}_3(\nu) = (\frac{\lambda p}{\nu + \lambda p})^3,$$

$$p_4(t) = \frac{(\lambda pt)^3}{3!} e^{-\lambda pt}.$$

І нарешті:

$$p_5(t) = 1 - (1 + \lambda pt + \frac{(\lambda pt)^2}{2!} + \frac{(\lambda pt)^3}{3!}) e^{-\lambda pt}.$$

При практичному застосуванні одержаних виразів необхідно задати $\lambda p$ кількість успішних нападів за одиницю часу. Величину інтенсивності потоку нападів $\lambda$ визначимо з аналізу статистичних даних, а імовірність $p$ подолання перешкоди визначимо через співвідношення ресурсів нападу і захисту $X$ і $Y$, або в статистичному плані – як частку $f$ вилученої з об'єкта інформації при заданому співвідношенні $x/y$. Цю залежність задамо у вигляді дрібно-лінійної функції:

$$\mathrm{f}(\mathrm{x},\mathrm{y}) = \frac{(\mathrm{x}/\mathrm{y})^2}{(\mathrm{x}/\mathrm{y})^2 + 16}.$$

Числові параметри в знаменнику цієї залежності визначаються з умови досягнення реальних, з нашої точки зору, величин $f$ при певних значеннях $x/y$ (наприклад, при $x/y=1$ $f(x,y)=0.10$; при $x/y=3$ $f(x,y)=0.21$; при $x/y \to \infty$ $f(x,y) \to 0.5$, а не $1$, що визначається початковою, в основному – природною захищеністю об'єкта).

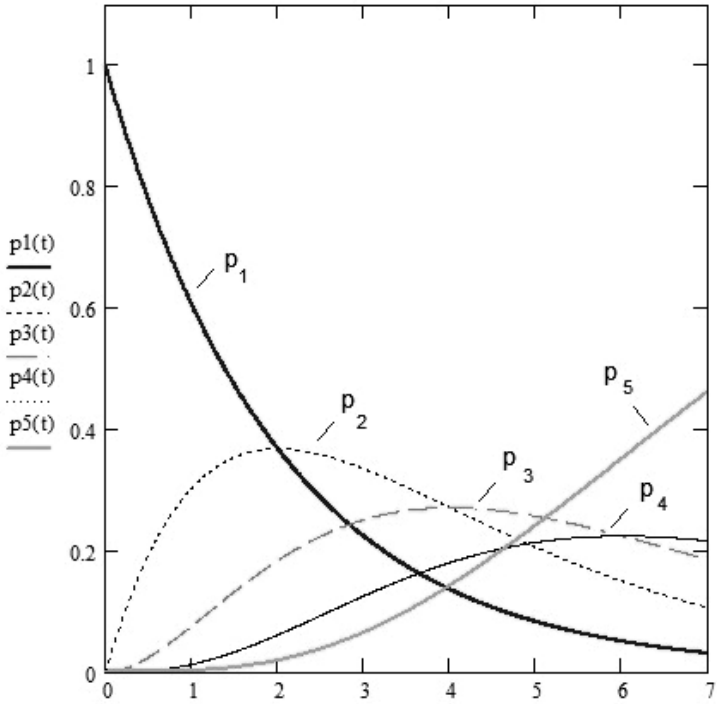


Рис. 13.9.Імовірності станів в залежності від часу.

*Приклад 6*. Візьмемо *λ=5/міс* (5 нападів на місяць); *x/y=1*, звідки *p=f(x,y)=0.1*. При цьому *λp=0.5 1/міс* (імовірність подолання однієї перешкоди за 1 місяць складає 50%). Імовірності станів, визначені приведеними виразами, зображені на рис. 13.9.

Значення $t_{im}$, при яких досягаються максимальні значення $p_{im}$ ймовірностей $p_i$, можна знайти з виразів для $p_i$. При $i = \overline{2,4}$ $t_{im} = \dfrac{i-1}{\lambda p}$. Величина $p_1(t)$ досягає максимуму $p_{1max} = 1$ при $t_{1max} = 0$, для $p_5(t)$ максимальне значення $p_{5max} = 1$ досягається при $t_{5max} \to \infty$.

Приведена методика носить ілюстративний характер і може бути розповсюджена на складніші системи, які відрізняються такими показниками:

1) кількість об'єктів і розташування перешкод;
2) співвідношення ресурсів нападу і захисту $\frac{X}{Y}$ ;
3) кількість інформації $g_k$ на об'єктах і форма залежності *f(x,y)* вразливості від величин *x* та *y* на об'єктах;
4) розподіл $x_k$ по об'єктах.

Подальший розвиток методики може вестись як в напрямку розрахунку додаткових показників (наприклад, кінцевого стану інформаційної системи), так і в бік ускладнення умов, зокрема комплексного протистояння.

## Контрольні питання

1. Основні положення теорії марковських ланцюгів. Дискретні і неперервні ланцюги.
2. Розрахунок імовірності станів інформаційної безпеки за даними розміченого графа.
3. Розрахунок перехідних ймовірностей на основі залежності *f(x)*.
4. Застосування неперервних марковських ланцюгів для визначення станів інформаційної безпеки.

## ЛІТЕРАТУРА


1. Беллман Р. Динамическое программирование. – М.: ИИЛ,1960. – 400с.

2. Бугір М. К. Математика для економістів. – К.: Академія, 2003. – 520 с.

3. Васильченко І. П. Вища математика для економістів. – К.: Знання, 2007. – 454 с.

4. Вентцель Е. С. Исследование операций. – М.: Сов.радио, 1972. – 552 с.

5. Гермейер Ю. Б. Введение в исследование операций. – М.: Наука, 1971. – 383 с.

6. Глухов В. В., Медников М.Д., Коробко С. Б. Математические методы и модели для менеджмента. – Спб.: Лань, 2005. – 528 с.

7. Дослідження операцій в економіці / За редакцією І. К. Федоренко, О.І. Черняка. – К.: Знання, 2007. – 558 с.

8. Задірака В.К., Олексюк О. С., Смоленюк Р. П., Штабалюк П.І.Фінансування витрат на захист інформації в економічній діяльності // «Університетські наукові записки». – 2006. - № 3-4 (19-20). - С. 479-490.

9. Івченко І. Ю. Математичне програмування. – К.: ЦУЛ, 2007. – 232 с.

10. Исследование операций в экономике / Под редакцией Н. Ш. Кремера. – М.: Юнити, 2006. – 407 с.

11. Карагодова О. О., Кігель В. Р., Рожок В. Д. Дослідження операцій. – К.: ЦУЛ, 2007. – 256 с.

12. Лабскер Л. Г., Бабешко Л. О. Игровые методы в управлении экономикой и бизнесом. – М.: Дело, 2001. – 464 с.

13. Левченко Є.Г., Рабчун А.О. Оптимізаційні задачі менеджменту інформаційної безпеки // Сучасний захист інформації. – 2010.– №1. – С. 16-23.

14. Применение теории игр в военном деле / Под редакцией В.О. Ашкеназы. – М.: Сов. радио, 1961. – 360 с.

15. Ржевський С. В., Александрова В.М. Дослідження операцій. – К.: Академвидав, 2006. – 558 с.

16. Саати Т. Принятие решений. Метод анализа иерархий: Пер. с англ.- М.: Радио и связь, 1993. – 278 с.

17. Таха Х. А. Введение в исследование операций. М.: Вильямс, 2005. – 901 с.

18. Тихонов В. И., Миронов М. А. Марковские процессы – М.: Сов. радио, 1977. – 488с.

19. Шикин Е.В., Шикина Г.Е. Исследование операций. – М.: Проспект, 2006. –280 с.

20. Bellman R.E., Zadeh L.A. Decision-making in a fuzzy environment // Management Science. – 1970. – Vol.17. – №4. – pp. 141 – 164.

21. Bohme R. Moor T. The iterated weakest link: A model of adaptive security investment // The Eights Worlshop of the Economics of Information Security. – University College London, UK. - 2009. - pp. 1 - 29.

22. Gordon L., Loeb M. The Economics of Information Security Investment. ACM Transactions of Information and System Security. – 2002. – Vol.5. –№4. – pp.438-457.

23. Liu W., Tanaka H., Matsuura K. – Empirical Analysis Methodology for Information Security Investment and Its Application to Reliable Survey of Japanese Firms //IPSJ Digital Courier. – 2006. – Vol.3. – pp. 585-599.

24. 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. – Hanover, USA. - 2008. - pp. 25 – 28.

25. Tatsumi K., Goto M. Optimal timing of information security investment // The Eights Worlshop of the Economics of Information Security. – University College London, UK. - 2009. - pp.

26. Zadeh L. A. Fuzzy sets // Information and Control. – 1965. – Vol. 8. – pp. 338 – 353.

27. 2006 CSI/FBI Computer Crime and Security Survey.

28. 2006 Japan and US Computer Crime and Security Survey.

29. 2006 Australian Computer Crime and Security Survey.

30. 2007 CSI/FBI Computer Crime and Security Survey.

31. 2008 CSI/FBI Computer Crime and Security Survey.

Навчальне видання

ЛЕВЧЕНКО Євген Григорович
ШВЕЦЬ Валеріян Анатолійович
ДЕМЧИШИН Мирослав Володимирович

**ЕКОНОМІКА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ**

Навчальний посібник

Укладачі: ЛЕВЧЕНКО Євген Григорович
ШВЕЦЬ Валеріян Анатолійович
ДЕМЧИШИН Мирослав Володимирович

Технічний редактор
Коректор
Комп'ютерна верстка

Підп. до друку . Формат 60х84/16. Папір офс.
Офс. друк. Ум. друк. арк. 14,125. Обл.-вид. арк. 14.
Тираж 300 пр. Замовлення № . Вид. №

Видавництво НАУ
03680. Київ-58, проспект Космонавта Комарова, 1.

Свідоцтво про внесення до Державного реєстру ДК № 977 від 05.07.2002