Лекція 30.

Моделювання систем масового обслуговування

Великий клас систем, які складно вивчити аналітичними способами, але які добре вивчаються методами статистичного моделювання, зводиться до систем масового обслуговування (СМО).

У СМО мається на увазі, що є *типові шляхи* (канали обслуговування), через які в процесі обробки проходять *заявки*. Прийнято говорити, що заявки *обслуговуються* каналами. Канали можуть бути різними по призначенню, характеристикам, вони можуть сполучатися в різних комбінаціях; заявки можуть перебувати в чергах й очікувати обслуговування. Частина заявок може бути обслужена каналами, а частини можуть відмовити в цьому. Важливо, що заявки, з погляду системи, абстрактні: це те, що бажає обслужитися, тобто пройти певний шлях у системі. Канали є також абстракцією: це те, що обслуговує заявки.

Заявки можуть приходити нерівномірно, канали можуть обслуговувати різні заявки за різний час і так далі, кількість заявок завжди досить велико. Все це робить такі системи складними для вивчення й керування, і простежити всі причинно-наслідкові зв'язки в них не представляється можливим. Тому прийнято подання про те, що обслуговування в складних системах носить випадковий характер.

Прикладами СМО (див. табл. 30.1) можуть служити: автобусний маршрут і перевезення пасажирів; виробничий конвеєр по обробці деталей; ескадрилья, що влітає на чужу територію, літаків, що «обслуговується» зенітками ПВО; стовбур і ріжок автомата, які «обслуговують» патрони; електричні заряди, що переміщаються в деякому пристрої й т.д.

Таблиця 30.1.
Приклади систем масового обслуговування

СМО	**Заявки**	**Канали**
Автобусний маршрут і перевезення пасажирів	Пасажири	Автобуси
Виробничий конвеєр по обробці деталей	Деталі, вузли	Верстати, склади
Ворожа територію для літаків, яка «обслуговується» зенітками ПВО	Літаки	Зенітні знаряддя, радари, стрілки, снаряди
Стовбур і ріжок автомата, які «обслуговують» патрони	Патрони	Стовбур, ріжок
Електричні заряди, що переміщаються в деякому пристрої	Заряди	Каскади технічного пристрою

Але всі ці системи об'єднані в один клас СМО, оскільки підхід до їхнього вивчення єдиний. Він полягає в тому, що, по-перше, за допомогою генератора випадкових чисел розігруються випадкові числа, які імітують ВИПАДКОВІ моменти появи заявок і час їхнього обслуговування в каналах. Але в сукупності ці випадкові числа, звичайно, підлеглі *статистичним* закономірностям.

Приміром, нехай сказано: «заявки в середньому приходять у кількості 5 штук у годину». Це означає, що часи між приходом двох сусідніх заявок випадкові, наприклад: 0.1; 0.3; 0.1; 0.4; 0.2, як це показано на мал. 30.1, але в сумі вони дають у середньому 1 (зверніть увагу, що в прикладі це не точно 1, а 1.1 — але зате в іншу годину ця сума, наприклад, може бути рівної 0.9); і тільки *за досить великий час* середнє цих чисел стане близьким до однієї години.

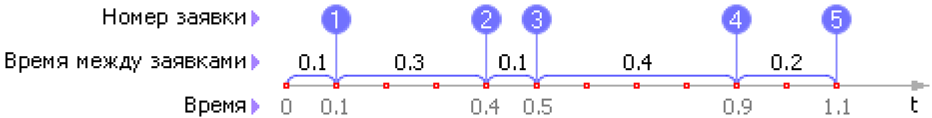


Рис. 30.1. Випадковий процес приходу заявок у СМО

Результат (наприклад, пропускна здатність системи), звичайно, теж буде випадковою величиною на окремих проміжках часу. Але обмірювана на великому проміжку часу, ця величина буде вже, у середньому, відповідати точному рішенню. Тобто для характеристики СМО цікавляться відповідями в статистичному змісті.

Отже, систему випробовують випадковими вхідними сигналами, підлеглими заданому статистичному закону, а як результат приймають статистичні показники, усереднені за часом розгляди або по кількості досвідів. Раніше, у лекції 21 (див. рис. 21.1), ми вже розробили схему для такого статистичного експерименту (див. мал. 30.2).

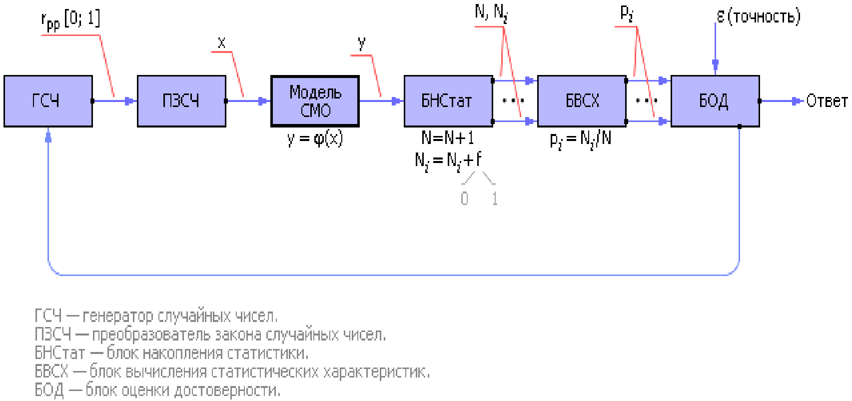


Рис. 30.2. Схема стат. експ-ту для вивчення систем масового обслуговування

По-друге, всі моделі СМО збираються типовим образом з невеликого набору елементів (канал, джерело заявок, черга, заявка, дисципліна обслуговування, стек, кільце й так далі), що дозволяє імітувати ці задачі *типовим* образом. Для цього модель системи збирають із конструктора таких елементів. Неважливо, яка конкретно система вивчається, важливо, що схема системи збирається з тих самих елементів. Зрозуміло, структура схеми буде завжди різною.

Перелічимо деякі основні поняття СМО.

Канали — те, що обслуговує; бувають гарячі (починають обслуговувати заявку в момент її надходження в канал) і холодні (каналу для початку обслуговування потрібен час на підготовку). **Джерела заявок** — породжують заявки у випадкові моменти часу, відповідно до заданого користувачем статистичному закону. **Заявки**, вони ж клієнти, входять у систему (породжуються джерелами заявок), проходять через її елементи (обслуговуються), залишають її обслуженими або незадоволеними. Бувають нетерплячі заявки — такі, котрим набридло очікувати або перебувати в системі і які залишають по власній волі СМО. Заявки утворять потоки — потік заявок на вході системи, потік обслужених заявок, потік відмовлених заявок. Потік характеризується кількістю заявок певного сорту, спостережуваним у деякому місці СМО за одиницю часу (година, доба, місяць), тобто потік є величина статистична.

Черги характеризуються правилами стояння в черзі (дисципліною обслуговування), кількістю місць у черзі (скільки клієнтів максимум може перебувати в черзі), структурою черги (зв'язок між місцями в черзі). Бувають обмежені й необмежені черги. Перелічимо найважливіші дисципліни обслуговування. FIFO (First In, First Out — першим прийшов, першим пішов): якщо заявка першої прийшла в чергу, то вона першої піде на обслуговування. LIFO (Last In, First Out — останнім прийшов, першим пішов): якщо заявка останньої прийшла в чергу, то вона першої піде на обслуговування (приклад — патрони в ріжку автомата). SF (Short Forward — короткі вперед): у першу чергу обслуговуються ті заявки із черги, які мають менший час обслуговування.

Дамо яскравий приклад, що показує, як правильний вибір тієї або іншої дисципліни обслуговування дозволяє одержати відчутну економію за часом.

Нехай є два магазини. У магазині № 1 обслуговування здійснюється по черзі, тобто тут реалізована дисципліна обслуговування FIFO (див. мал. 30.3).

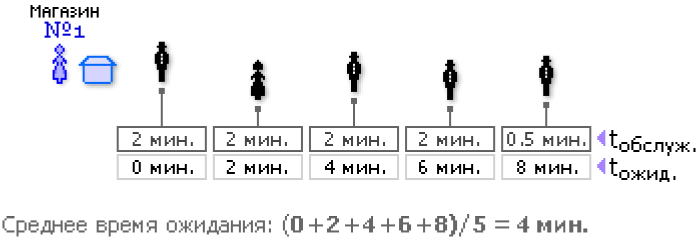


Рис. 30.3. Організація черги по дисципліні FIFO

Час обслуговування $t_{обслуж.}$ на мал. 30.3 показує, скільки часу продавець затратить на обслуговування одного покупця. Зрозуміло, що при покупці штучного товару продавець затратить менше часу на обслуговування, чим при покупці, скажемо, сипучих продуктів, що вимагають додаткових маніпуляцій (набрати, зважити, вирахувати ціну й т. п). Час очікування $t_{ожид.}$ показує, через який час черговий покупець буде обслужений продавцем.

У магазині № 2 реалізована дисципліна SF (див. мал. 30.4), що означає, що штучний товар можна купити позачергово, тому що час обслуговування $t_{обслуж.}$ такої покупки невеликий.

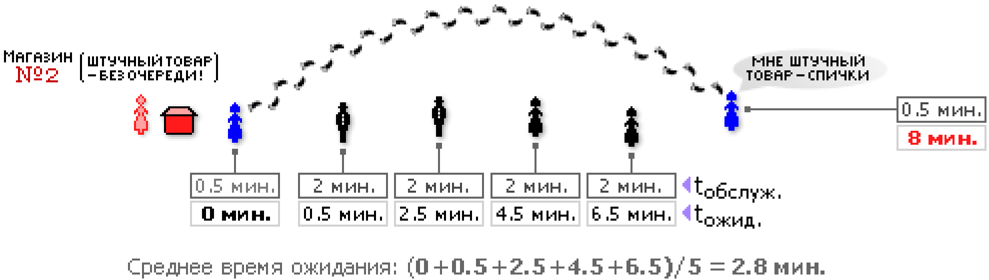


Рис. 30.4. Організація черги по дисципліні SF

Як видно з обох малюнків, останній (п'ятий) покупець збирається придбати штучний товар, тому час його обслуговування невеликий — 0.5 хвилин. Якщо цей покупець прийде в магазин № 1, він буде змушений вистояти в черзі цілих 8 хвилин, у той час як у магазині № 2 його обслужать відразу ж, позачергово. Таким чином, середній час обслуговування кожного з покупців у магазині з дисципліною обслуговування FIFO складе 4 хвилини, а в магазині з дисципліною обслуговування КВ — лише 2.8 хвилини. А суспільна користь, економія часу складе: $(1 – 2.8/4) \cdot 100\% = 30$ відсотків! Отже, 30% зекономленого для суспільства часу - і це лише за рахунок правильного вибору дисципліни обслуговування.

Фахівець із моделювання складних систем повинен добре розуміти ресурси продуктивності й ефективності проектованих їм систем, сховані в оптимізації параметрів, структур і дисциплінах обслуговування. Моделювання допомагає виявити ці сховані резерви.

При аналізі результатів моделювання важливо також указати інтереси й ступінь їхнього виконання. Розрізняють інтереси клієнта й інтереси власника системи. Помітимо, що ці інтереси збігаються не завжди.

Судити про результати роботи СМО можна по показниках. Найбільш популярні з них:

- імовірність обслуговування клієнта системою;
- пропускна здатність системи;
- імовірність відмови клієнтові в обслуговуванні;
- імовірність зайнятості кожного з каналу й всіх разом;
- середній час зайнятості кожного каналу;
- імовірність зайнятості всіх каналів;
- середня кількість зайнятих каналів;
- імовірність простою кожного каналу;
- імовірність простою всієї системи;
- середня кількість заявок, що коштують у черзі;
- середній час очікування заявки в черзі;
- середній час обслуговування заявки;
- середній час знаходження заявки в системі.

Судити про якість отриманої системи потрібно по сукупності значень показників. При аналізі результатів моделювання (показників) важливо також звертати увагу на інтереси клієнта й інтереси власника системи, тобто мінімізувати або максимізувати треба той або інший показник, а також на ступінь їхнього виконання. Помітимо, що найчастіше інтереси клієнта й власника між собою не збігаються або збігаються не завжди. Показники будемо позначати далі $H = \{h_1, h_2, \ldots\}$...

Параметрами СМО можуть бути: інтенсивність потоку заявок, інтенсивність потоку обслуговування, середній час, протягом якого заявка готова очікувати обслуговування в черги, кількість каналів обслуговування, дисципліна обслуговування й так далі. Параметри — це те, що впливає на показники системи. Параметри будемо позначати далі як $R = \{r_1, r_2, \ldots\}$...

Приклад. Автозаправна станція (АЗС).

1. Постановка задачі. На мал. 30.5 наведений план АЗС. Розглянемо метод моделювання СМО на її прикладі й план її дослідження. Водії, проїжджаючи по дорозі мимо АЗС по дорозі, можуть захотіти заправити свій автомобіль. Хочуть обслуговуватися (заправити машину бензином) не всі автомобілісти підряд; допустимо, що із усього потоку машин на заправлення в середньому заїжджає 5 машин у годину.

Рис. 30.5. План модельованої АЗС

На АЗС два однакові стовпчики, статистична продуктивність кожної з яких відома. Перший стовпчик у середньому обслуговує 1 машину в годину, друга в середньому — 3 машини в годину. Власник АЗС заасфальтував для машин місце, де вони можуть очікувати обслуговування. Якщо стовпчики зайняті, то на цьому місці можуть очікувати обслуговування інші машини, але не більше двох одночасно. Черга будемо вважати загальної. Як тільки один зі стовпчиків звільниться, то перша машина із черги може зайняти її місце на колонку (при цьому друга машина просувається на перше місце в черзі). Якщо з'являється третя машина, а всі місця (їх два) у черзі зайняті, то їй відмовляють в обслуговуванні, тому що стояти на дорозі забороненої (див. дорожні знаки біля АЗС). Така машина їде ладь із системи назавжди і як потенційний клієнт є загубленою для власника АЗС. Можна ускладнити задачу, розглянувши касу (ще один канал обслуговування, куди треба потрапити після обслуговування в одній зі стовпчиків) і черга до неї й так далі.

Але в найпростішому варіанті очевидно, що шляхи руху потоків заявок по СМО можна зобразити у вигляді еквівалентної схеми, а додавши значення й позначення характеристик кожного елемента СМО, одержуємо остаточно схему, зображену на мал. 30.6.

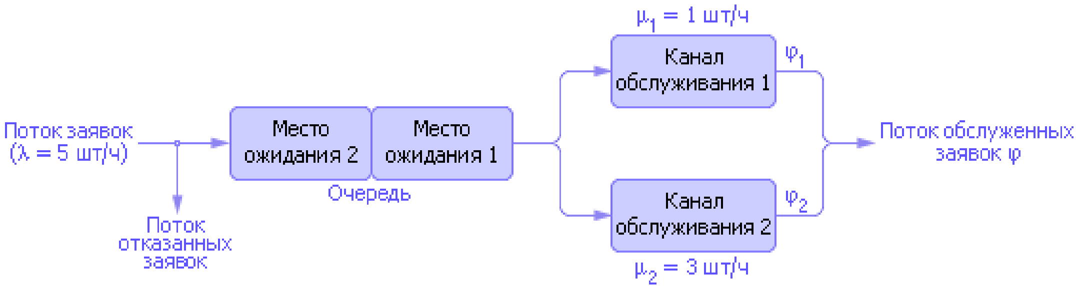


Рис. 30.6. Еквівалентна схема об'єкта моделювання

2. Метод дослідження СМО. Застосуємо в нашому прикладі принцип послідовної проводки заявок (докладно про принципи моделювання див. лекцію 32). Його ідея полягає в тім, що заявку проводять через всю систему від входу до виходу, і тільки після цього беруться за моделювання наступної заявки.

Для наочності побудуємо часову діаграму роботи СМО, відбиваючи на кожній лінійці (вісь часу t) стан окремого елемента системи. Часових лінійок проводиться стільки, скільки є різних місць у СМО, потоків. У нашому прикладі їх 7 (потік заявок, потік очікування на першому місці в черги, потік очікування на другому місці в черги, потік обслуговування в каналі 1, потік обслуговування в каналі 2, потік обслужених системою заявок, потік відмовлених заявок).

Для генерації часу приходу заявок використаємо формулу обчислення інтервалу між моментами приходу двох випадкових подій (див. лекцію 28):

$$t = -\frac{1}{\lambda} \cdot \ln(r)$$

У цій формулі величина потоку λ повинна бути задана (до цього вона повинна бути визначена експериментально на об'єкті як статистичне середнє), r — випадкове рівномірно розподілене число від 0 до 1 із ГВЧ або таблиці, у якій випадкові числа потрібно брати підряд (не вибираючи спеціально).

Задача. Згенеруйте потік з 10 випадкових подій з інтенсивністю появи подій 5 шт/година.

Рішення задачі. Візьмемо випадкові числа, рівномірно розподілені в інтервалі від 0 до 1 (див. таблицю), і обчислимо їхні натуральні логарифми (див. табл. 30.2).

Таблиця 30.2.

Фрагмент таблиці випадкових чисел й їхніх логарифмів

r_{pp}[0; 1]	ln(r_{pp}[0; 1])
0. 0333	–3. 4022
0. 3557	–1. 0337
0. 2172	–1. 5269
0. 5370	–0. 6218

Формула пуасонівського потоку визначає відстань між двома випадковими подіями в такий спосіб: $t = -\mathrm{Ln}(r_{pp})/\lambda$. Тоді, з огляду на, що $\lambda = 5$, маємо відстані між двома випадковими сусідніми подіями: 0.68, 0.21, 0.31, 0.12 години. Тобто події наступають: перше — у момент часу $t = 0$, друге — у момент часу $t = 0.68$, третє — у момент часу $t = 0.89$, четверте — у момент часу $t = 1.20$, п'яте — у момент часу $t = 1.32$ і так далі. Події — прихід заявок відіб'ємо на першій лінійці (див. мал. 30.7).

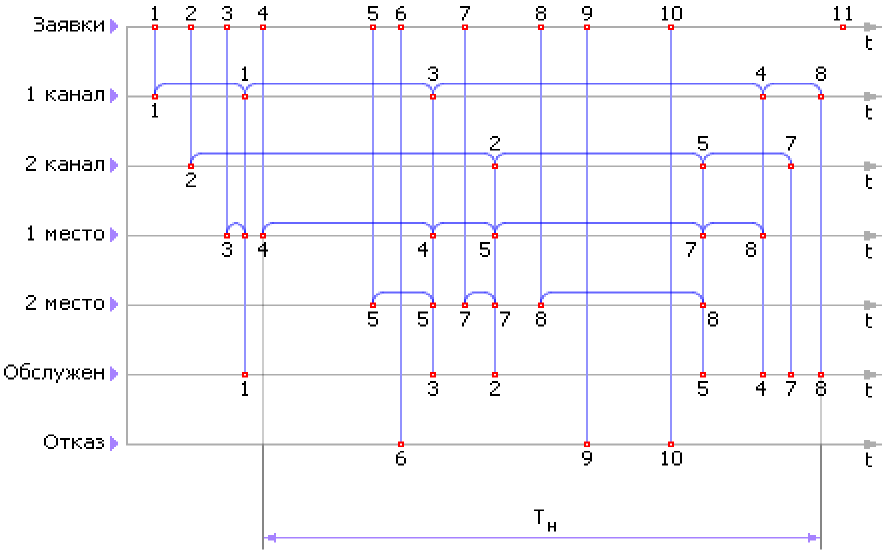


Рис. 30.7. Часова діаграма роботи СМО

Береться перша заявка й, тому що в цей момент канали вільні, установлюється на обслуговування в перший канал. Заявка 1 переноситься на лінійку «1 канал».

Час обслуговування в каналі теж випадкове й обчислюється по аналогічній формулі:

$$t = -\frac{1}{\lambda} \cdot \ln(r)$$

де роль інтенсивності грає величина потоку обслуговування μ_1 або μ_2, залежно від того, який канал обслуговує заявку. Знаходимо на діаграмі момент закінчення обслуговування, відкладаючи згенерований час обслуговування від моменту початку обслуговування, і опускаємо заявку на лінійку «Обслужені».

Заявка пройшла в СМО весь шлях. Тепер можна, відповідно до принципу послідовної проводки заявок, також проімітувати шлях другої заявки.

Якщо в деякий момент виявиться, що обидва канали зайняті, то варто встановити заявку в чергу. На мал. 30.7 це заявка з номером 3. Помітимо, що за умовами задачі в черзі на відміну від каналів заявки перебувають не випадковий час, а очікують, коли звільниться якийсь із каналів. Після звільнення каналу заявка піднімається на лінійку відповідного каналу й там організується її обслуговування.

Якщо всі місця в черзі в момент, коли прийде чергова заявка, будуть зайняті, то заявку варто відправити на лінійку «Відмовлені». На мал. 30.7 це заявка з номером 6.

Процедуру імітації обслуговування заявок продовжують якийсь час спостереження $T_н$. Чим більше цей час, тим точніше надалі будуть результати моделювання. Реально для простих систем вибирають $T_н$, рівне 50—100 і більше годин, хоча іноді краще міряти цю величину кількістю розглянутих заявок.

Аналіз часової діаграми

Аналіз проведемо на вже розглянутому прикладі.

Спочатку потрібно дочекатися сталого режиму. Відкидаємо перші чотири заявки як нехарактерні, що протікають під час процесу встановлення роботи системи. Вимірюємо час спостереження, допустимо, що в нашому прикладі воно складе $T_н = 5$ годин. Підраховуємо з діаграми кількість обслужених заявок $N_{обс.}$, часи простою й інші величини. У результаті можемо обчислити показники, що характеризують якість роботи СМО.

1. Імовірність обслуговування: $P_{обс.} = N_{обс.}/N = 5/7 = 0.714$. Для розрахунку ймовірності обслуговування заявки в системі досить розділити число заявок, яким удалося обслужитися за час $T_н$ (див. лінійку «Обслужені») $N_{обс.}$, на число заявок N, які хотіли обслужитися за цей же час. Як і раніше ймовірність експериментально визначаємо відношенням подій, що відбулись, до загального числа подій, які могли відбутися!

2. Пропускна здатність системи: $A = N_{обс.}/T_{н} = 7/5 = 1.4$ [шт/година]. Для розрахунку пропускної здатності системи досить розділити число обслужених заявок $N_{обс.}$ на час $T_{н}$, за яке відбулося це обслуговування (див. лінійку «Обслужені»).
3. Імовірність відмови: $P_{отк.} = N_{отк.}/N = 3/7 = 0.43$. Для розрахунку ймовірності відмови заявці в обслуговуванні досить розділити число заявок $N_{отк.}$, яким відмовили за час $T_{н}$ (див. лінійку «Відмовлені»), на число заявок N, які хотіли обслужитися за цей же час, тобто надійшли в систему. **Зверніть увагу**. $P_{отк.} + P_{обс.}$ у теорії повинне бути дорівнює 1. Насправді експериментально вийшло, що $P_{отк.} + P_{обс.} = 0.714 + 0.43 = 1.144$. Ця неточність порозумівається тим, що час спостереження $T_{н}$ мало й статистика накопичена недостатня для одержання точної відповіді. Погрішність це показника зараз становить 14%!

4. Вероятность зайнятості одного каналу: $P_1 = T_{\text{зан.}}/T_{\text{н}} = 0.05/5 = 0.01$, де $T_{\text{зан.}}$ — час зайнятості тільки одного каналу (першого або другого). Вимірам підлягають часові відрізки, на яких відбуваються певні події. Наприклад, на діаграмі шукаються такі відрізки, під час яких зайняті або перший або другий канал. У даному прикладі є один такий відрізок наприкінці діаграми довжиною 0.05 години. Частка цього відрізка в загальному часі розгляду ($T_{\text{н}} = 5$ годин) визначається діленням і становить шукану ймовірність зайнятості.
5. Імовірність зайнятості двох каналів: $P_2 = T_{\text{зан.}}/T_{\text{н}} = 4.95/5 = 0.99$. На діаграмі шукаються такі відрізки, під час яких одночасно зайняті й перший, і другий канал. У даному прикладі таких відрізків чотири, їхня сума дорівнює 4.95 години. Частка тривалості ці події в загальному часі розгляду ($T_{\text{н}} = 5$ годин) визначається діленням і становить шукану ймовірність зайнятості.

6. Середня кількість зайнятих каналів:
   $N_{\text{ск}} = 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 = 0.01 + 2 \cdot 0.99 = 1.99$. Щоб підрахувати, скільки каналів зайнято в системі в середньому, досить знати частку (імовірність зайнятості одного каналу) і помножити на вагу цієї частки (один канал), знати частку (імовірність зайнятості двох каналів) і помножити на вагу цієї частки (два канали) і так далі. Отримана цифра 1.99 говорить про те, що з можливих двох каналів у середньому завантажено 1.99 каналу. Це високий показник завантаження, 99.5%, система добре використає ресурс.
7. Імовірність простою хоча б одного каналу:
   $P^*{}_1 = T_{\text{простоя}1}/T_{\text{н}} = 0.05/5 = 0.01$.
8. Імовірність простою двох каналів одночасно: $P^*{}_2 = T_{\text{простоя}2}/T_{\text{н}} = 0$.
9. Імовірність простою всієї системи: $P^*{}_{\text{с}} = T_{\text{простоя сист.}}/T_{\text{н}} = 0$.

10. Середня кількість заявок у черзі:
$N_{сз} = 0 \cdot P_{0з} + 1 \cdot P_{1з} + 2 \cdot P_{2з} = 0.34 + 2 \cdot 0.64 = 1.62$ [шт]. Щоб визначити середня кількість заявок у черзі, треба визначити окремо ймовірність того, що в черзі буде одна заявка $P_{1з}$, імовірність того, у черзі буде стояти дві заявки $P_{2з}$ и т. буд. і знову з відповідними вагами їх скласти.

11. Імовірність того, що в черзі буде одна заявка:
$P_{1з} = T_{1з}/T_{н} = 1.7/5 = 0.34$ (усього на діаграмі чотирьох таких відрізка, у сумі дающих 1.7 години).

12. Імовірність того, у черзі буде стояти одночасно дві заявки:
$P_{2з} = T_{2з}/T_{н} = 3.2/5 = 0.64$ (усього на діаграмі три таких відрізки, у сумі дающих 3.25 години).

13. Середній час очікування заявки в черзі:

$$T_{ср.ож.} = \frac{\sum_{i=1}^{N} T_{ож.i}}{N}$$

(Скласти всі часові інтервали, протягом яких яка-небудь заявка перебувала в черзі, і розділити на кількість заявок). На часовій діаграмі таких заявок 4.

14. Середній час обслуговування заявки:

$$T_{ср.обсл.} = \frac{\sum_{i=1}^{N} T_{обсл.i}}{N}$$

(Скласти всі часові інтервали, протягом яких яка-небудь заявка перебувала на обслуговуванні в якому-небудь каналі, і розділити на кількість заявок).

15. Середній час знаходження заявки в системі:

$T_{\text{ср. сист.}} = T_{\text{ср. ож.}} + T_{\text{ср. обсл.}}$.

16. Середня кількість заявок у системі:

$$N_{cp.} = \frac{\sum_{i=1}^{K}(N_{2i} + N_{3i} + N_{4i} + N_{5i})}{K}$$

Розіб'ємо інтервал спостереження, наприклад, на десятихвилинки. Вийде на п'ятьох годинниках K підінтервалів (у нашому випадку $K = 30$). У кожному підінтервалі визначимо по часовій діаграмі, скільки заявок у цей момент перебуває в системі. Дивитися треба на 2, 3, 4 й 5-ю лінійки — які з них зайняті в цей момент. Потім суму K *що* складають усереднити.

Далі варто оцінити точність кожного з отриманих результатів. Тобто відповістити на запитання: наскільки ми можемо довіряти цим значенням? Оцінка точності проводиться за методикою, описаної в лекції 34.

Якщо точність не є задовільної, то варто збільшити час експерименту й тим самим поліпшити статистику. Можна зробити й по-іншому. Снову кілька разів запустити експеримент на час $T_н$. А в наслідку усереднити значення цих експериментів. І знову перевірити результати на критерій точності. Цю процедуру варто повторювати доти, поки не буде досягнута необхідна точність.

Далі варто скласти таблицю результатів й оцінити значення кожного з них з погляду клієнта й власника СМО (див. табл. 30.3).. Наприкінці, з огляду на сказане в кожному пункті, варто зробити загальний висновок. Таблиця повинна мати приблизно такий вид, який показаний у табл. 30.3.

Таблиця 30.3.
Показники СМО

Показник	**Формула**	**Значення**	**Інтереси власника СМО**	**Інтереси клієнта СМО**
Імовірність обслуговування	$P_{\text{обс.}} = N_{\text{обс.}}/N$	0. 714	Імовірність обслуговування мала, багато клієнтів ідуть із системи незадоволеними, їхні гроші для власника загублені. Це «мінус». Рекомендація: збільшити ймовірність обслуговування.	Імовірність обслуговування мала, кожен третій клієнт хоче, але не може обслужитися. Це «мінус». Рекомендація: збільшити ймовірність обслуговування.
…	…	…	…	…

Середня кількість заявок у черзі	$N_{сз} = 0 \cdot P_{0з} + 1 \cdot P_{1з} + 2 \cdot P_{2з}$	1.62	Черга практично увесь час вся забита. Всі місця в черзі використаються досить ефективно. Вкладення на організацію черги окупають витрати на неї. Це «плюс». Клієнти, які довго коштують у черзі, можуть піти, не дочекавшись обслуговування. Клієнти, простоюючи, можуть завдати шкоди системі,	Черга практично увесь час вся забита. Клієнтові доводиться стояти в черзі, перш ніж він потрапить на обслуговування. Клієнт може не потрапити навіть у чергу. Це «мінус». Рекомендація: збільшити пропускну здатність,

			ламати встаткування. Багато відмов, загублених клієнтів. Це «мінуси». Рекомендація: збільшити число місць у черзі, збільшити пропускну здатність.	збільшити число місць у черзі.

Загальний підсумок:	В інтересах власника: а) збільшити пропускну здатність каналів, щоб не губити клієнтів (правда, модернізація каналів коштує грошей); б) збільшити число місць у черзі (це теж коштує грошей), щоб затримати потенційних клієнтів.	Клієнти зацікавлені в значному збільшенні пропускної здатності для зменшення часу очікування й зменшення відмов.

Синтез СМО

Ми пробили аналіз існуючої системи. Це дало можливість побачити її недоліки й визначити напрямку поліпшення її якості. Але залишаються незрозумілими відповіді на конкретні питання, що саме треба зробити - збільшувати кількість каналів або збільшувати їхню пропускну здатність, або збільшувати кількість місць у черзі, і, якщо збільшувати, то наскільки? Є й такі питання, що краще - створити 3 канали із продуктивністю 5 прим/годину або один із продуктивністю 15 шт/година?

Щоб оцінити чутливість кожного показника до зміни значення певного параметра, надходять у такий спосіб. Фіксують всі параметри крім одного, обраного. Потім знімають значення всіх показників при декількох значеннях цього обраного параметра. Звичайно, доводиться повторювати знову й знову процедуру імітації й усреднять показники при кожнім значенні параметра, оцінювати точність. Але в результаті виходять надійні статистичні залежності характеристик (показників) від параметра.

Наприклад, для 12 показників нашого приклада можна одержати 12 залежностей від одного параметра: залежність імовірності відмов $P_{\text{отк.}}$ від кількості місць у черзі (СМО), залежність пропускної здатності A від кількості місць у черзі, і так далі (див. мал. 30.8).

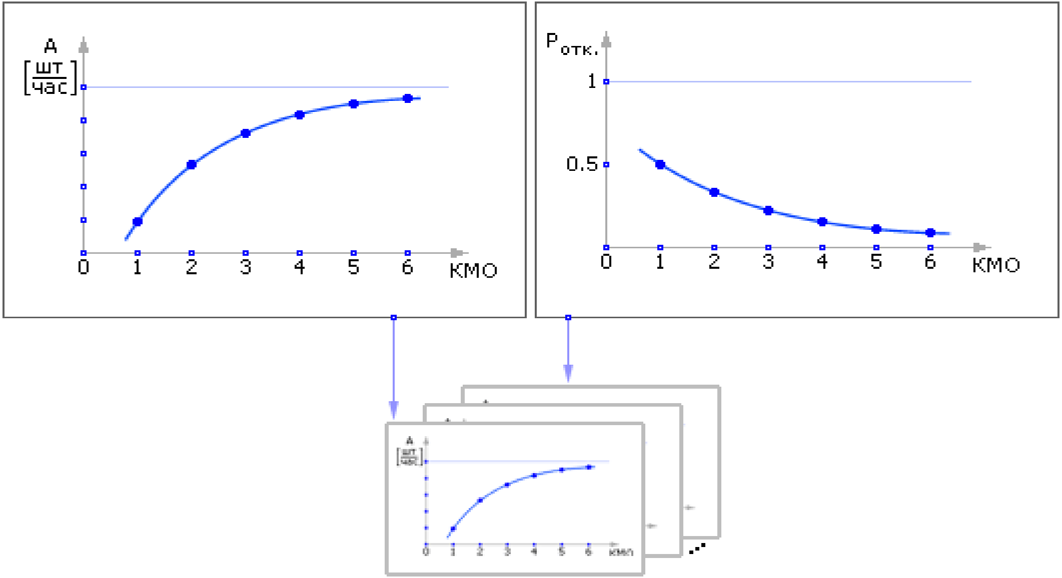


Рис. 30.8. Зразковий вид залежностей показників від параметрів СМО

Потім так само можна зняти ще 12 залежностей показників P від іншого параметра R, зафіксувавши інші параметри. І так далі. Утвориться своєрідна матриця залежностей показників P від параметрів R, по якій можна провести додатковий аналіз про перспективи руху (поліпшення показників) у ту або іншу сторону. Нахил кривих добре показує чутливість, ефект від руху по певному показнику. У математику цю матрицю називають якобіаном J, у якій роль нахилу кривих грають значення похідних $\Delta P_i/\Delta R_j$, див. мал. 30.9. (Нагадаємо, що похідна зв'язана геометрично з кутом нахилу дотичній до залежності.)

$$\|J\| = \begin{vmatrix} J_{11} & J_{12} & J_{13} & \cdots & J_{1,12} \\ J_{21} & J_{22} & J_{23} & \cdots & J_{2,12} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ J_{51} & J_{52} & J_{53} & \cdots & J_{5,12} \end{vmatrix} = \begin{vmatrix} \dfrac{\Delta P_1}{\Delta R_1} & \dfrac{\Delta P_1}{\Delta R_2} & \dfrac{\Delta P_1}{\Delta R_3} & \cdots & \dfrac{\Delta P_1}{\Delta R_{12}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \dfrac{\Delta P_5}{\Delta R_1} & \dfrac{\Delta P_5}{\Delta R_2} & \dfrac{\Delta P_5}{\Delta R_3} & \cdots & \dfrac{\Delta P_5}{\Delta R_{12}} \end{vmatrix} =$$

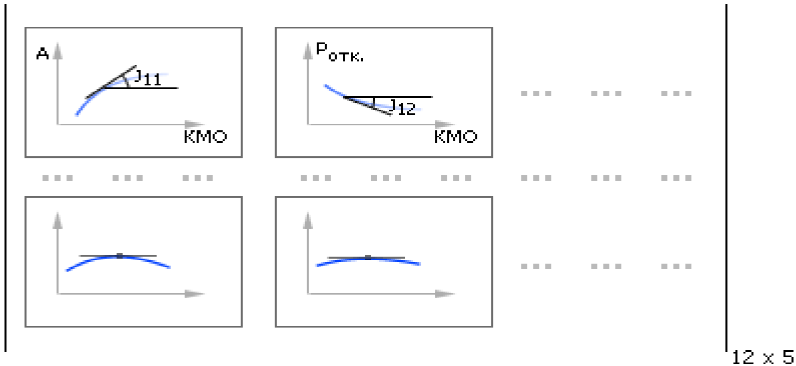
$=$ 

Рис. 30.9. Якобіан — матриця чутливостей показників залежно від зміни параметрів СМО

Якщо показників 12, а параметрів, наприклад, 5, то матриця має розмірність 12 x 5. Кожен елемент матриці — крива, залежність *i-го* показника від *j-го* параметра. Кожна крапка кривої — середнє значення показника на досить представницькому відрізку $T_н$ або усереднено по декількох експериментах.

Варто розуміти, що криві знімалися в припущенні того, що всі параметри крім одного в процесі їхнього зняття були незмінні. (Якби всі параметри міняли значення, те криві були б іншими. Але так не роблять, тому що вийде повна плутанина й залежностей не буде видно.)

Тому, якщо на підставі розгляду знятих кривих приймається рішення про те, що деякий параметр буде в СМО змінений, те всі криві для нової крапки, у якій знову буде досліджуватися питання про те, який параметр варто змінити, щоб поліпшити показники, варто знімати заново.

Так крок за кроком можна спробувати поліпшити якість системи. Але поки ця методика не може відповісти на ряд питань. Справа в тому, що, по-перше, якщо криві монотонно ростуть, то виникає питання, де ж все-таки варто зупинитися. По-друге, можуть виникати протиріччя, один показник може поліпшуватися при зміні обраного параметра, у той час як іншої буде одночасно погіршуватися. По-третє, ряд параметрів складно виразити чисельно, наприклад, зміна дисципліни обслуговування, зміна напрямків потоків, зміна топології СМО. Пошук рішення у двох останніх випадках проводиться із застосуванням методів експертизи (див. лекцію 36) і методами штучного інтелекту.

Тому зараз обговоримо тільки перше питання. Як прийняти рішення, яким повинне бути все-таки значення параметра, якщо з його ростом показник увесь час монотонно поліпшується? Навряд чи значення нескінченності влаштує інженера.

Параметр R — керування, це те, що перебуває в розпорядженні власника СМО (наприклад, можливість заасфальтувати площадку й тим самим збільшити кількість місць у черзі, поставити додаткові канали, збільшити потік заявок за рахунок збільшення витрат на рекламу й так далі). Міняючи керування, можна впливати на значення показника P, ціль, критерій (імовірність відмов, пропускну здатність, середній час обслуговування й так далі). З мал. 30.10 видно, що якщо збільшувати керування R, те можна домогтися завжди поліпшення показника P. Але очевидно, що будь-яке керування пов'язане з витратами Z. І чим більше докладають зусиль для керування, чим більше значення керуючого параметра, тим більше витрати. Звичайно витрати на керування ростуть лінійно: $Z = C_1 \cdot R$. Хоча зустрічаються випадки, коли, наприклад, в ієрархічних системах, вони ростуть експоненційно, іноді - обернено експоненційно (знижки за опт) і так далі.

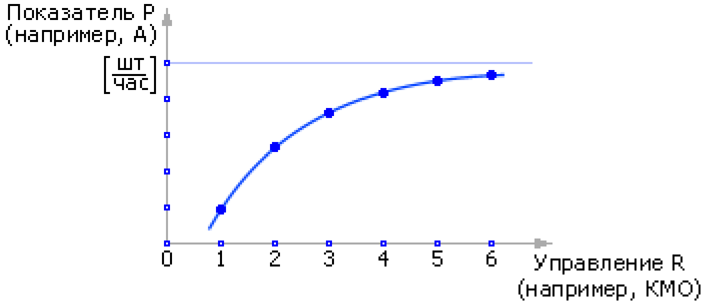


Рис. 30.10. Залежність показника P
від керованого параметра R (приклад)

У кожному разі ясно, що колись вкладення всі нових витрат просто перестане себе окупати. Наприклад, ефект від заасфальтованої площадки розміром в 1 км^2 навряд чи окупить витрати власника бензоколонки в Задрипінську, там просто не буде стільки бажаючих заправитися бензином. Іншими словами показник *P* у складних системах не може рости нескінченно. Рано або пізно його ріст сповільнюється. А витрати *Z* ростуть (див. мал. 30.11).

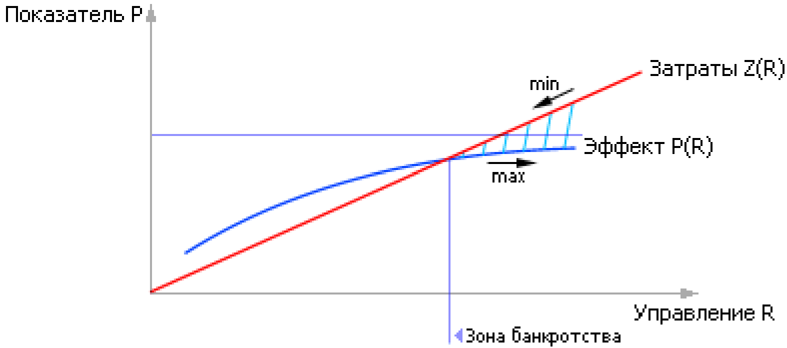


Рис. 30.11. Залежності ефекту від застосування показника P і витрат Z на його одержання як функції керованого параметра R

З мал. 30.11 видно, що при призначенні ціни C_1 за одиницю витрат R і ціни C_2 за одиницю показника P, ці криві можна скласти. Криві складають, якщо їх потрібно одночасно мінімізувати або максимізувати. Якщо одна крива підлягає максимізації, а інша мінімізації, то варто знайти їхню різницю, наприклад по крапках. Тоді результуюча крива (див. мал. 30.12), що враховує й ефект від керування й витрати на це, буде мати екстремум. Значення параметра R, що доставляє екстремум функції, і є рішення задачі синтезу.

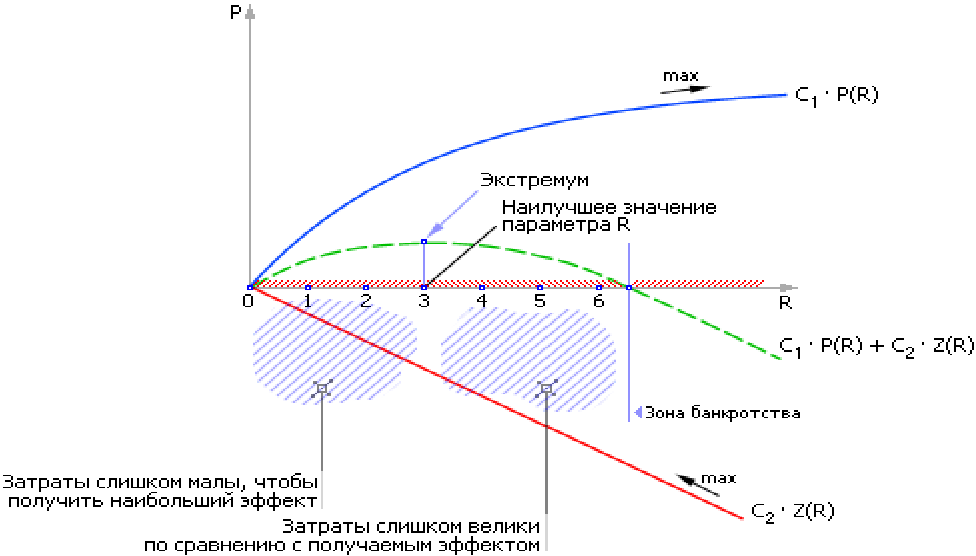


Рис. 30.12. Сумарна залежність ефекту від застосування показника P
и витрат Z на його одержання як функції керованого параметра R

Крім керування R і показника P у системах діє збурювання. Збурювання позначимо як $D = \{d_1, d_2, \ldots\}$, см. мал. 30.13. Збурювання — це вхідний вплив, що, на відміну від керуючого параметра, не залежить від волі власника системи. Наприклад, низькі температури на вулиці, конкуренція знижують, на жаль, потік клієнтів, поломки встаткування прикро знижують продуктивність системи. І управляти цими величинами безпосередньо власник системи не може. Звичайне збурювання діє «на зло» власникові, знижуючи ефект P від керуючих зусиль R. Це відбувається тому, що, у загальному випадку, система створюється для досягнення цілей, недосяжних самих по собі в природі. Людина, організувати систему, завжди сподівається за допомогою її досягти деякої мети *P*. На це він затрачає зусилля *R*, ідучи всупереч природі. Система — організація доступних людині, вивчених їм природних компонентів для досягнення деякої нової мети, недосяжної раніше іншими способами.

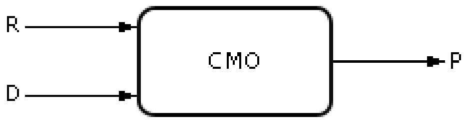


Рис. 30.13. Умовна позначка досліджуваної системи, на яку впливають керуючі впливи R і збурювання D

Отже, якщо ми знімемо залежність показника *P* від керування *R* ще раз (як показаний на мал. 30.10), але в умовах збурювання, що *з'явився*, D, те, можливо, характер кривої зміниться. Швидше за все, показник буде при однакових значеннях керувань перебувати нижче, тому що збурювання носить «противний» характер, знижуючи показники системи (див. мал. 30.14). **Система, надана сама собі, без зусиль керуючого характеру, перестає забезпечувати мета, для досягнення якої вона була створена**. Якщо, як і раніше, побудувати залежність витрат, співвіднести її із залежністю показника від параметра керування, то знайдена крапка экстремума зміститься (див. мал. 30.15) у порівнянні з випадком «збурювання = 0» (див. мал. 30.12).

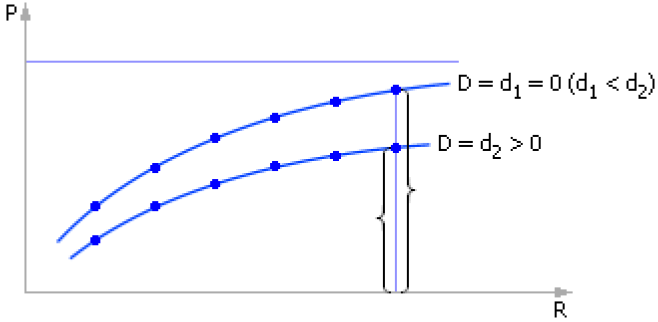


Рис. 30.14. Залежність показника P від керуючого параметра R при різних значеннях діючих на систему збурювань D

Якщо знову збільшити збурювання, то криві зміняться (див. мал. 30.14) і, як наслідок, знову зміниться положення крапки екстремуму (див. мал. 30.15).

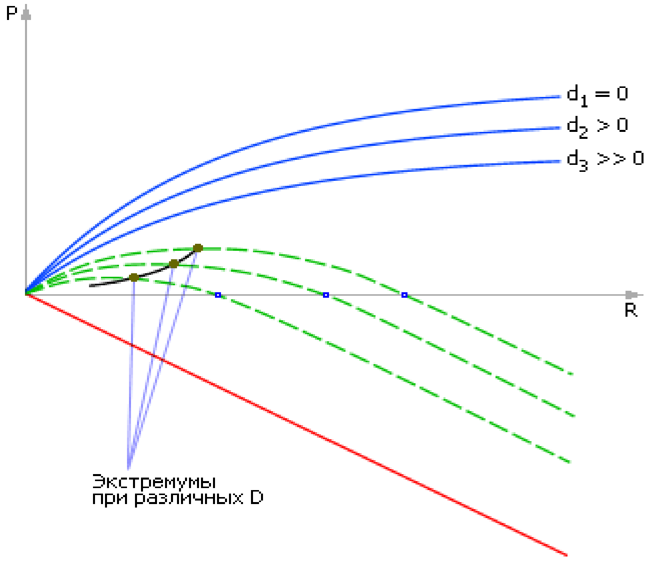


Рис. 30.15. Знаходження крапки екстремуму на сумарній залежності при різних значеннях діючого фактору, що обурює, D

В остаточному підсумку, всі знайдені положення крапок екстремуму переносяться на новий графік, де утворять залежність *Показника P* від *Керуючого параметра R* при зміні *Збурювань D* (див. мал. 30.16).

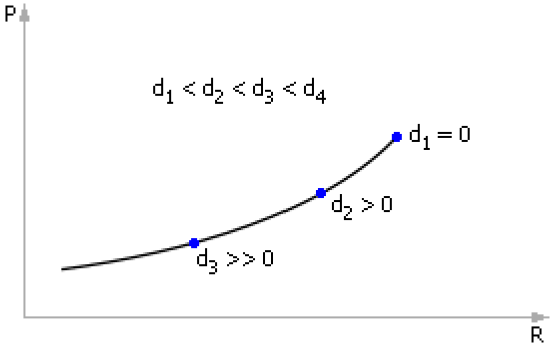


Рис. 30.16. Залежність показника P від керуючого параметра R при зміні значень збурювань D (крива складається тільки із крапок екстремумів)

Зверніть увагу, що насправді на цьому графіку можуть бути й інші робочі крапки (графік пронизаний як би сімействами кривих), але нанесені нами крапки задають такі координати керуючого параметра, при яких при заданих збурюваннях (!) досягається найбільше з можливих значення показника P.

Цей графік (див. мал. 30.16) зв'язує Показник P, Керування (ресурс) R і Збурювання D у складних системах, указуючи, як діяти щонайкраще ОПР (особі, що приймає рішення) в умовах виниклих збурювань. Тепер користувач може, знаючи реальну обстановку на об'єкті (значення збурювання), швидко за графіком визначити, яке керуючий вплив на об'єкт необхідно, щоб забезпечити найкраще значення його показника, що цікавить.

Помітьте, якщо керуючий вплив буде менше оптимального, те сумарний ефект знизиться, виникне ситуація недоотриманого прибутку. Якщо керуючий вплив буде більше оптимального, то ефект **також** знизиться, тому що заплатити за чергове збільшення керуючих зусиль треба буде по величині більше, ніж та, котру ви одержите в результаті її використання (ситуація банкрутства).

Примітка. У тексті лекції ми використали слова «керування» й «ресурс», тобто вважали, що $R = U$. Варто пояснити, що керування дійсно відіграє роль деякої обмеженої цінності для власника системи. Тобто завжди є коштовним для нього ресурсом, за який завжди доводиться платити, і якого завжди не вистачає. Дійсно, якби ця величина не була обмежена, то ми б могли досягати за рахунки нескінченної величини керувань нескінченно більших значень цілей, а от нескінченно більших результатів у природі явно не спостерігається.

Іноді розрізняють властиво керування U і ресурс R, називаючи ресурсом деякий запас, тобто границю можливого значення керуючого впливу. У цьому випадку поняття ресурс і керування не збігаються: $U < R$. Іноді розрізняють граничне значення керування $U \leq R$ й інтегральний ресурс $\int U dt \leq R$.