

С. Б. Дубіненко
Ю. І. Мітюшкін

ДОСЛІДЖЕННЯ ОПЕРАЦІЙ

Ч. 1. СИСТЕМИ МАСОВОГО ОБСЛУГОВУВАННЯ

Міністерство освіти і науки України
Вінницький національний технічний університет

С. Б. Дубіненко
Ю. І. Мітюшкін

ДОСЛІДЖЕННЯ ОПЕРАЦІЙ

Ч. 1. СИСТЕМИ МАСОВОГО ОБСЛУГОВУВАННЯ

Затверджено Вченою радою Вінницького національного технічного університету як навчальний посібник для студентів напрямку підготовки “Системна інженерія” всіх спеціальностей. Протокол № 7 від 27 грудня 2007 р.

Вінниця ВНТУ 2008

УДК 519.872 (075)

Д 79

Рецензенти:

С. В. Юхимчук, доктор технічних наук, професор

В. М. Лисогор, доктор технічних наук, професор

В. М. Папінов, кандидат технічних наук, доцент

Рекомендовано до видання Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України

Дубіненко С. Б., Мітюшкін Ю. І.

Д 79 Дослідження операцій. Ч. 1. Системи масового обслуговування.
Навчальний посібник. – Вінниця: ВНТУ, 2008. – 55 с.

В першій частині посібника узагальнюються математичні основи моделювання систем масового обслуговування. На основі базових понять та властивостей детально розглядаються марковські моделі систем масового обслуговування, а також основні типи немарковських моделей. Посібник розроблений у відповідності з планом кафедри та програмами дисциплін “Математичне програмування та дослідження операцій”, “Планування експерименту”, “Теорія надійності”.

УДК 519.872 (075)

ЗМІСТ

Вступ.....	4
1 Загальні відомості про системи масового обслуговування.....	6
1.1 Класифікація систем масового обслуговування (СМО)	6
1.2 Основні проблеми, які розв'язуються методами ТМО.....	9
1.3 Показники ефективності СМО.....	10
1.4 Характеристики потоків вимог і обслуговуючих каналів.....	12
1.4.1 Потоки вимог.....	12
1.4.2 Найпростіший потік і його властивості.....	15
1.4.3 Розподіл інтервалів часу між двома подіями найпростішого потoku.	20
1.4.4 Основні характеристики обслуговуючих приладів.....	22
Питання для самоконтролю.....	24
2 Марковські моделі систем масового обслуговування.....	25
2.1 Система $M / M / 1 / \infty$	25
2.1.1 Диференціальний метод аналізу (рівняння станів)	26
2.1.2 Визначення ймовірнісних і моментних характеристик СМО при сталому режимі.....	28
2.1.3 Інтегральний метод аналізу СМО.....	30
2.1.4 Вихідний потік вимог.....	32
2.2 Багатоканальні системи масового обслуговування ($M/M/S/\infty$).....	34
2.2.1 Багатоканальна СМО з однаковими приладами.....	34
2.2.2 Багатоканальна СМО з неоднаковими приладами.....	37
2.3 Замкнені СМО.....	38
2.4 Багатофазні СМО.....	40
2.5 Основні дисципліни обслуговування вимог вхідних потоків.....	42
2.5.1 Загальні положення.....	42
2.5.2 Вплив дисципліни з відносним пріоритетом на час очікування в черзі.....	43
Питання для самоконтролю.....	45
3 Немарковські моделі масового обслуговування.....	46
3.1 Система $M / G / 1 / \infty$	46
3.1.1 Система $M / E / 1 / \infty$	46
3.1.2 Система $M / D / 1 / \infty$	48
3.2 Обслуговування неоднорідних потоків вимог.....	48
3.3 Необхідність врахування неоднорідності потоків вимог, що надходять до системи.....	52
Питання для самоконтролю.....	53
Література	54

ВСТУП

Важливою проблемою нашого суспільства є інтенсивне впровадження систем підтримки прийняття рішень (СППР), яке неможливе без відповідної теорії прийняття рішень. Важливим розділом цієї теорії є дослідження операцій (ДО). В дослідженні операцій особлива увага приділяється методам прийняття рішень в різних ситуаціях: в умовах повної невизначеності, при наявності випадкових факторів, невизначеності в цілях тощо.

Як самостійний науковий напрямок, ДО сформувалося на початку 40-х років минулого століття. Перші публікації були присвячені розв'язанню воєнних задач, зокрема, аналізу та дослідженню бойових операцій. Звідти й виникла назва дисципліни. Великий внесок у формування та розвиток ДО внесли Л. В. Кондратович, Б. В. Гнеденко, Н. П. Бусленко, Д. Б. Юдін, Н. П. Федоренко, Р. Акоф, Р. Бекман, Г. Кун, Т. Сааті, А. Кофман та інші.

Дослідження операцій – це наука, яка займається розробленням та практичним застосуванням методів оптимального управління організаційними системами.

Предмет ДО можна поділити на три частини:

- 1) детерміновані моделі;
- 2) ймовірнісні моделі;
- 3) нелінійні моделі.

Навчальний посібник має на меті викладення основ теорії масового обслуговування (ТМО), яка є складовою частиною ймовірнісних моделей. При написанні посібника автори використали матеріали робіт [1-7], які рекомендовані студентам для більш детального вивчення.

В загальному випадку будь-яку систему, в якій потік вимог зустрічає обмежені засоби їх задоволення, можна розглядати як систему масового обслуговування (СМО). До таких систем, включаючи засоби і канали зв'язку, відносяться ЕОМ, телемеханічні пристрої, верстати, механізми, транспортні засоби (повітряні, наземні, морські), засоби і підприємства обслуговування (магазини, автобуси, лікарні тощо), виробничі процеси (експлуатація пристроїв і механізмів, складальні лінії, обслуговування верстатів тощо), керування запасами, фізичні процеси (наприклад, рух частинок) та багато інших.

На сьогоднішній день методи і результати ТМО знаходять широке застосування при вирішенні проблем теорії надійності, аналізу процесів функціонування складних систем, при розробленні різних АСУ та інших систем, до яких надходять потоки вимог різної фізичної природи. ТМО дозволяє формалізувати різні процеси керування і системи, незалежно від їх конкретного змісту. Це пояснюється єдиною термінологією описання процесів, що відбуваються в системі: потік вимог, обслуговування вимог, обслуговуючий прилад (канал), дисципліна обслуговування. При цьому тео-

рія ставить задачу – достатньо повно описати сутність явищ, що відбуваються в системі, і встановити з достатньою для практики точністю чисельний зв'язок між кількістю приладів обслуговування, характеристиками вхідного потоку і якістю обслуговування, під якою мається на увазі своєчасність проведення обслуговування вимог.

Під вхідним потоком вимог будемо розуміти впливи будь-якої природи, що надходять до системи через деякі випадкові проміжки часу. Ці вимоги система повинна обслужити. Сама назва “система масового обслуговування” свідчить про те, що надходження вимог носить масовий, а не поодинокий характер. Наприклад, потік завдань для ЕОМ, потік покупців до магазину, потік відмов апаратури ототожнюються з потоком вимог, що надходять до системи.

Обслуговування слід розуміти в самому широкому сенсі цього поняття: обслужити заявку чи вимогу означає, що необхідно затратити деяку кількість роботи, виконати певну кількість операцій, затратити деякий, зазвичай випадковий, час на перероблення і видозміну вимоги, що надійшла. Наприклад, якщо ця вимога являє собою відмову пристрою, то її обслуговування характеризується часом усунення цієї відмови.

В загальному випадку під терміном "обслуговування" будемо розуміти задоволення певних потреб, що оцінюється витратами часу або кількістю виконаних операцій. В процесі обслуговування вимог система переходить з одного стану в інший. Під обслуговуючим приладом (каналом) слід розуміти засіб або пристрій, що виконує обслуговування вимог, які надходять. При цьому процес обслуговування характеризує перехід приладу з одного стану в інший. Далі під приладом будемо розуміти також людину, якщо вона зайнята певним видом обслуговування. Обслуговані вимоги залишають прилад, утворюючи вихідний потік вимог.

Якщо моменти надходження вимог і (або) тривалість їх обслуговування не регламентуються, то при цьому можуть виникати конфлікти, що виражаються в утворенні черги вимог перед обслуговуючим приладом. Довжина цієї черги залежить від інтенсивності надходження вимог, її статистичної флуктуації і пропускну здатності обслуговуючого приладу. Коли інтенсивність потоку вимог перевищує пропускну здатність приладу, черга досягає необмеженої довжини. Однак навіть при інтенсивності надходження вимог, меншій за пропускну здатність приладу, черга може утворитися за рахунок статистичної флуктуації інтенсивності потоку і часу обслуговування вимог.

В загальному випадку під системою масового обслуговування будемо розуміти сукупність вхідного потоку (чи потоків), обслуговуючого приладу (приладів) і черги, що може характеризуватися кількістю вимог, які в ній знаходяться. Довжина черги визначає затримку в обслуговуванні вимог, що надходять, і може використовуватись при оцінюванні якості роботи системи масового обслуговування. ТМО дозволяє оцінити вплив випадкових факторів і характеристик системи на утворення черги. Через ва-

жливість такого показника якості роботи системи, як утворення черги, в англomовній літературі теорію масового обслуговування називають ще теорією черг.

1 ЗАГАЛЬНІ ВІДОМОСТІ ПРО СИСТЕМИ МАСОВОГО ОБСЛУГОВУВАННЯ

1.1 Класифікація систем масового обслуговування (СМО)

Будь-яка система характеризується своєю структурою, тобто складом її основних частин і функціональних зв'язків між ними. Для СМО такими частинами є обслуговуючі прилади (канали).

Якщо система містить тільки один канал, така СМО називається одноканальною. Однак у деяких випадках система може містити кілька каналів, призначених для одночасного (паралельного) обслуговування вимог, що надходять. Такі СМО називають багатоканальними. Прикладами зазначених СМО є кілька ЕОМ, що працюють паралельно, кілька каналів зв'язку, по яких одночасно передається інформація (повідомлення), кілька продавців у магазині тощо. Графічна модель таких СМО показана на рис. 1.1.

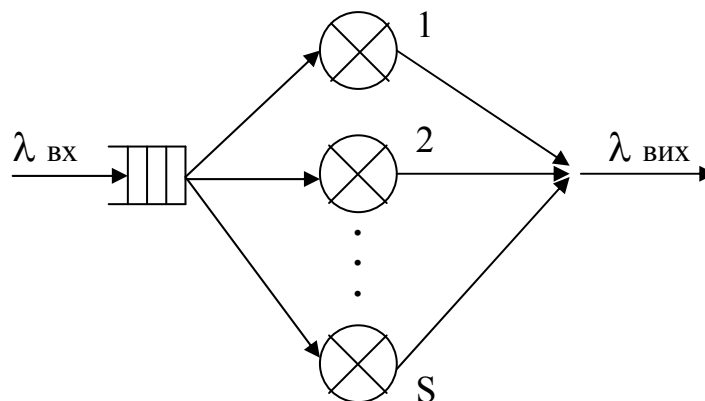


Рисунок 1.1 – Графічна модель багатоканальних СМО

Спеціалізація окремих каналів по певних операціях приводить до багатофазової СМО, кожна фаза якої може містити в собі один чи кілька каналів, що обслуговують той самий (для фази) потік вимог. У цьому випадку вимога вважається повністю виконаною тільки тоді, коли вона послідовно пройде усі фази обслуговування. Прикладами багатофазних СМО є ЕОМ з її складовими частинами, виробничий конвеєр, виготовлення виробів на кількох ділянках і організація контролю їх якості, магазин, у якому покупців спочатку обслуговують продавці, а потім касир та інші. Графічна модель багатофазової СМО зображена на рис.1.2, з якого видно, що фазу можна розглядати як одно- чи багатоканальну СМО.

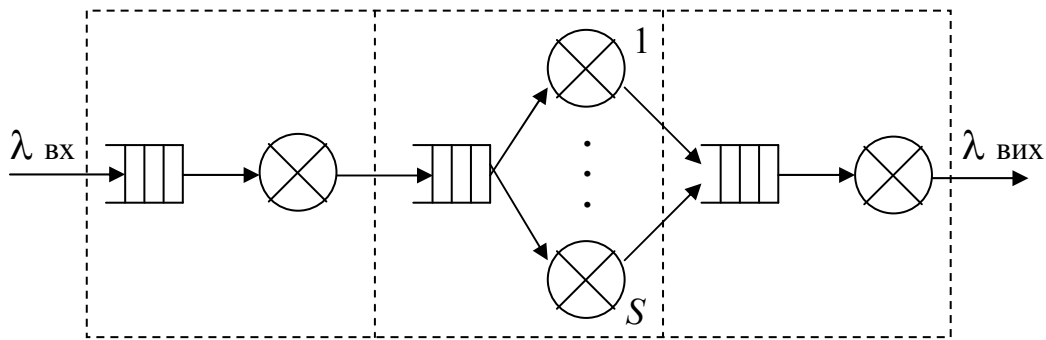


Рисунок 1.2 - Графічна модель багатофазової СМО

Закони розподілу потоків вимог і часу їх обслуговування істотно впливають на значення показників якості роботи системи. В практиці проектування та експлуатації реальних систем найчастіше зустрічаються такі закони розподілу: пуасонівський – M_1 , експоненціальний – M , регулярний – D , ерлангенський – E , довільний – G . Символи означають умовні скорочення назв цих законів.

В залежності від закону розподілу потоків вимог і часу їх обслуговування розрізняють пуасонівські і непуасонівські СМО, які описуються власне марківськими і немарківськими моделями масового обслуговування. Для характеристики СМО використовується символіка, запропонована Д. Кендаллом, що містить у собі п'ять розрядів, призначених для характеристики: потоку вимог (перший розряд), закону розподілу часу обслуговування вимог (другий розряд), особливості структури (одноканальна чи багатоканальна СМО – третій розряд), особливості черги (обмежена чи необмежена – четвертий розряд), а також для опису дисципліни обслуговування (п'ятий розряд). Останній розряд, зазвичай, опускається, тому що дисципліна обслуговування обумовлюється окремо.

Так, запис СМО $M / E_k / S / \infty$ означає: вхідний потік пуасонівський, час обслуговування вимог розподілено по ерлангенському закону k -го порядку, кількість обслуговуючих каналів – S , довжина черги перед приладами не обмежується. В окремих випадках можуть використовуватися і додаткові позначення.

В залежності від особливостей реальних систем і їх призначення, на довжину черги перед фазами СМО можуть накладатися обмеження. У цьому випадку розрізняють СМО з обмеженою і необмеженою довжиною черги. Вимоги, що надходять на СМО з обмеженою чергою, вважаються загубленими, якщо в момент їх надходження черга повністю заповнена. Окремим випадком СМО з обмеженою чергою є випадок, коли черга зовсім неприпустима. СМО, для яких вимоги можуть губитися, називаються СМО з втратами. Для СМО з необмеженою чергою вимоги, що надходять і застають усі канали зайнятими, очікують початку свого обслуговування в черзі доти, доки канал не звільниться від обслуговування вимог, які надійшли раніше, або більш пріоритетних.

Кількісні характеристики якості роботи СМО значною мірою залежать від установленого порядку (дисципліни) обслуговування вимог, що надходять. Деяким з них надається пріоритет в обслуговуванні. Наприклад, Герої Радянського Союзу обслуговуються в громадських організаціях поза чергою, термінові телеграми доставляються першими, незважаючи на те, що вони прийняті пізніше звичайних. Тому розрізняють СМО з пріоритетним обслуговуванням і СМО з дисципліною "першим прийшов – першим обслугований" (справедлива дисципліна). Один з різновидів графічної моделі з пріоритетним обслуговуванням показано на рис.1.3, де O_i – черга пріоритетних вимог ($i=1,2 \dots$); чим більше i , тим менше пріоритет.

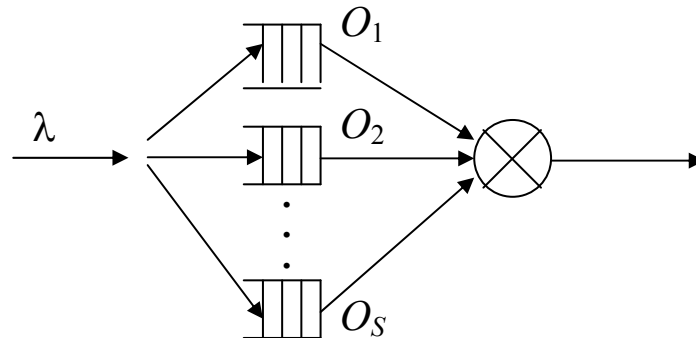


Рисунок 1.3 – Приклад графічної моделі з пріоритетним обслуговуванням

В залежності від загальної кількості джерел вимог U_i ($i=1,K$), кількості вимог, що направляються від них, і повернення обслугованих вимог до джерела розрізняють замкнені і розімкнені СМО.

Замкненими системами (системи з обмеженим вхідним потоком) називають системи, що призначені для обслуговування обмеженої кількості джерел вимог. Загальна кількість вимог, що циркулюють у замкнутій системі, скінченна (зазвичай постійна). Прикладом замкнутої системи може бути система, яка складається з декількох технічних пристроїв, що обслуговуються одним чи кількома операторами. Вихід з ладу пристрою є вимога (заявка) на обслуговування. Загальна кількість таких заявок не може бути більша числа закріплених за операторами пристроїв. Відновлений пристрій знову служить джерелом заявок. Схема замкнутої СМО показана на рис.1.4.

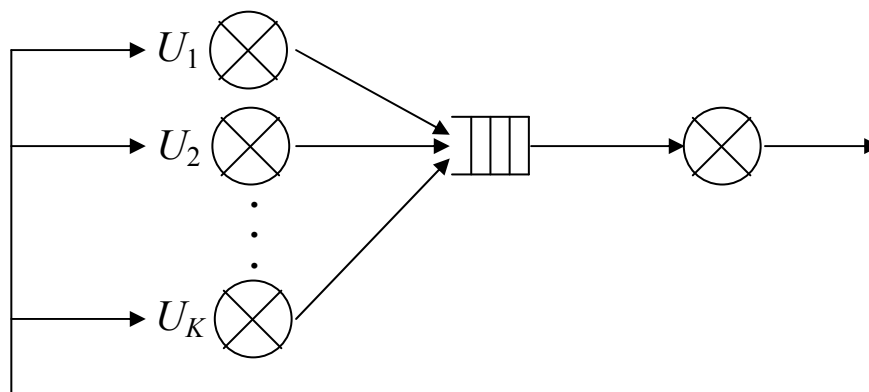


Рисунок 1.4 - Схема замкнутої СМО

У тому випадку, коли кількість джерел вимог необмежена або велика, або коли одне джерело посиляє в систему необмежену кількість вимог, СМО називають розімкненими. Прикладами розімкнених СМО служать ЕОМ, магазин, аеропорт, бригада робітників тощо.

Для кожної з розглянутих видів СМО розроблені відповідні математичні моделі, за допомогою яких можуть бути розраховані реальні системи, що їм відповідають. При цьому необхідно зазначити, що для більшості СМО вхідний потік передбачається однорідним та однотипним. На практиці зустрічаються потоки, що істотно відрізняються значеннями середнього часу обслуговування. Такі потоки називають неоднорідними. Дослідження СМО з неоднорідними потоками необхідні для розв'язання багатьох важливих задач практики проектування й експлуатації складних систем.

1.2 Основні проблеми, які розв'язуються методами ТМО

Методи ТМО за допомогою одних і тих самих математичних моделей при незначній їх зміні дозволяють розв'язувати такі задачі проектування, організації виробництва й експлуатації складних систем:

- розрахунок ефективності, надійності і потрібної кількості технічних пристроїв чи їх складових частин;
- прогнозування розвитку системи при зміні початкових умов (поява нових джерел вимог, зміна інтенсивності потоків тощо);
- планування різних робіт і операцій, необхідних для підвищення ефективності різних систем (виробництво, морські й авіаційні порти тощо);
- визначення раціональної кількості обслуговуючого персоналу, місць його розташування, необхідної кваліфікації тощо;
- керування запасами для виробництва складних систем і організації їх експлуатації;
- раціональна організація взаємодії складних систем чи їх компонентів;
- оптимізація процесів керування, оцінювання оптимальних параметрів систем, їх компонентів тощо.

Сама теорія масового обслуговування безпосередньо не пов'язана з оптимізацією. Вона скоріше намагається розробити, вивчити і порівняти різні ситуації і, таким чином, побічно досягти наближеної оптимізації.

Методи теорії масового обслуговування більшою мірою базуються на аналізі, аніж на синтезі. Методи синтезу використовуються для узагальнення, визначення напрямків аналітичних досліджень і вирішення проблем, що виникають. На відміну від теорії оптимізації, у якій головна задача полягає в знаходженні максимуму чи мінімуму цільової функції, основною рисою ТМО є математична формалізація досліджуваного процесу чи

системи. Теорія виражає у вигляді формул, пояснює і прогнозує ситуації масового обслуговування, забезпечуючи краще розуміння проблем, що виникають, і ухвалення раціонального рішення. Одержані за допомогою методів ТМО рішення будуть тим кращими, чим адекватніше в математичних моделях СМО відображаються реальні потоки і закони розподілу обслуговування вимог.

1.3 Показники ефективності СМО

Будь-яка система створюється для розв'язання певних задач. Тому необхідно мати такі показники якості роботи СМО, за допомогою яких можна судити про ефективність системи у виконанні задач, що стоять перед нею, чи покладених на неї функцій, і які дозволяють визначати оптимальні значення параметрів системи.

Розглянути всі можливі показники якості роботи СМО – задача нездійсненна, тому що не можна передбачити і перелічити усі способи практичного використання реальних систем.

До показників ефективності СМО відносять: імовірність того, що в черзі чи в системі знаходиться певна кількість вимог n ; середня кількість вимог у черзі чи в системі; розподіл часу очікування в черзі чи часу перебування вимог у системі; середній час і дисперсія часу очікування; імовірність того, що час очікування буде тривати не більше допустимого; імовірність обслуговування без очікування; середня тривалість інтервалу зайнятості системи; середня кількість обслугованих вимог; імовірність того, що обслуговування здійснюється певною кількістю каналів; середня кількість незайнятих каналів тощо.

У багатьох задачах істотну роль відіграє аналіз економічних витрат: вартість всієї операції, вартість обслуговуючих каналів, вартість збитку від несвоєчасного обслуговування, вартість розміщення черг різної довжини, вартість організації керування потоками вимог тощо. Розглянемо найбільш загальні й основні показники для різних реальних систем.

Усі показники (чи кількісні характеристики) якості обслуговування систем можна умовно підрозділити на три види: імовірнісні, моментні та економічні (вартісні), які тісно взаємопов'язані між собою. Серед імовірнісних характеристик широко застосовуються такі, як p_n – імовірність того, що в системі знаходиться n вимог; імовірність того, що вимога, яка надходить, буде загублена тощо.

З моментних характеристик до найбільш важливих відносяться:

- середня кількість вимог у системі: $\bar{n} = \sum_{n=0}^{\infty} n p_n$;

- середня кількість вимог у черзі: $\bar{v} = \sum_{n=S+1}^{\infty} (n - S) p_n$;

- середня кількість вимог, що обслуговуються, у фазі (системі):

$$\bar{j} = \sum_{n=0}^S np_n ;$$

- середня кількість вільних приладів: $\bar{r} = \sum_{n=0}^{\infty} (S - n)p_n ;$

- середній час очікування в черзі (нагромаджувачі): $\bar{\omega} = \int_0^{\infty} \omega dF(\omega) ,$ де

$F(x)$ – функція розподілу часу очікування в черзі;

- середня тривалість перебування вимоги в системі: $\bar{u} = \int_0^{\infty} u dG(u) ,$ де

$G(x)$ – імовірність того, що тривалість перебування в системі менше або дорівнює x .

Слід зазначити найбільш важливі й очевидні співвідношення між наведеними характеристиками. Для кожного моменту часу справедливі рівності:

$$\bar{j} + \bar{r} = \bar{S} , \quad \bar{\gamma} + \bar{j} = \bar{n} , \quad \bar{t}_{об} + \bar{\omega} = \bar{u} ,$$

де $\bar{t}_{об}$ – середня тривалість обслуговування вимог.

До цих загальних співвідношень можна додати інші. Зокрема, середній вхідний потік дорівнює середньому вихідному потоку: $\lambda_{вх} = \lambda_{вих} = \lambda$, коли СМО працює в сталому режимі роботи. З цієї властивості у випадку, коли вхідні вимоги не залежать від вихідних (розімкнена система), одержуємо рівності:

$$\lambda \bar{\omega} = \bar{\gamma} , \quad \lambda \bar{u} = \bar{n} .$$

Очевидно, що чим менше значення таких показників, як $\bar{\omega}$, $\bar{\gamma}$, $\bar{n}$, $\bar{u}$, тим вища ефективність системи з точки зору джерел вимог. Однак зменшення значень цих показників пов'язане з витратами на підвищення продуктивності і надійності обслуговуючих каналів. Конфлікт, що виникає при цьому між підвищенням продуктивності СМО і витратами, може бути вирішений за допомогою економічних характеристик.

Економічні характеристики якості обслуговування найчастіше подають таким узагальненим критерієм, що враховує збиток, викликаний простоєм приладів і витратами через очікування вимог. Функції вартості й обмеження при цьому виражаються за допомогою таких характеристик:

- єдиних цін чи інших величин (наприклад, вартість очікування в розрахунку на одну вимогу за одиницю часу і вартість простою в розрахунку на канал за одиницю часу);

- числових характеристик досліджуваного явища, що розглянуті раніше.

До складу узагальненого критерію включають вартість очікування вимог в черзі і вартість простою каналів. Тоді, задавши інтервал часу T , для якого потрібно розрахувати повну середню вартість $\Gamma(S)$, і позначивши як C_1 вартість очікування однієї вимоги за одиницю часу, а C_2 – вартість простою одного каналу за одиницю часу, одержимо:

$$\Gamma(S) = (C_1 \bar{\gamma} + C_2 \bar{r})T$$

або

$$\Gamma(S) = (C_1 \sum_{n=S+1}^{\infty} (n-S)p_n + C_2 \sum_{n=0}^S (S-n)p_n)T. \quad (1.1)$$

Змінною величиною в (1.1) найчастіше буває кількість каналів S , хоча змінною може бути також інтенсивність обслуговування μ каналу або інтенсивність вхідного потоку λ . Якщо вартості C_1 і C_2 постійні на інтервалі часу T , то легко знайти повну вартість за одиницю часу:

$$\gamma(S) = \frac{\Gamma(S)}{T} = (C_1 \sum_{n=S+1}^{\infty} (n-S)p_n + C_2 \sum_{n=0}^S (S-n)p_n).$$

Нерідко намагаються мінімізувати саме цю функцію, в результаті чого визначають оптимальну кількість каналів для системи. Зазначимо, що мінімальне значення $\gamma(S)$ рідко визначають аналітичними методами; зазвичай його знаходять чисельними методами.

1.4 Характеристики потоків вимог і обслуговуючих каналів

Для побудови математичних моделей СМО, знаходження значень показників ефективності необхідно знати особливості потоків, що надходять, і обслуговуючих каналів. В першу чергу це стосується законів розподілу потоків вимог і часу їх обслуговування.

1.4.1 Потоки вимог

Потоком вимог $\varphi(t)$ називається послідовність появи вимог (одна за одною) на вході системи. Якщо позначити моменти їх появи t_i ($i = 1, 2, 3, \dots$ – номер вимоги, що надійшла), то випадковий процес $\varphi(t)$ може бути поданий у вигляді окремих реалізацій, показаних на рис.1.5.

Нехай інтервали $\theta_1, \theta_2, \dots$ між появами вимог (див. рис. 1.5, а) є незалежними між собою. Тоді такий потік вимог називають потоком Пальма, чи потоком з обмеженою післядією. На рис. 1.5, б показана реалізація $\varphi(t)$, яка характеризує кількість вимог $N(t)$, що надійшли за час t .

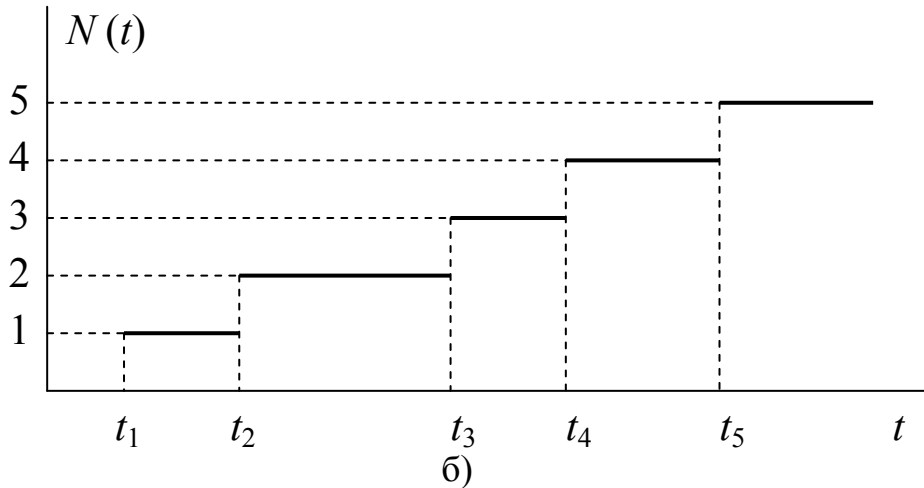
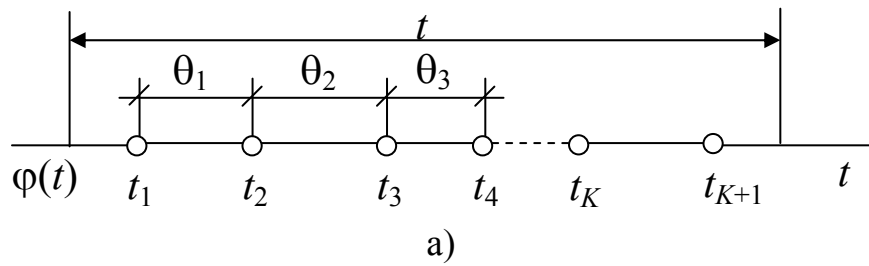


Рисунок 1.5 – Варіанти реалізації випадкового процесу $\varphi(t)$

Як інтервали θ_i , так і кількість вимог $N(t)$ в загальному випадку являють собою випадкові величини, які описуються деяким законом розподілу (функцією щільності або розподілом імовірності подій).

Найбільш простим з точки зору описання є регулярний потік, вимоги якого надходять на вхід системи через суворо однакові (регулярні) інтервали часу θ , що описуються функцією розподілу

$$F(t) = \begin{cases} 0, & t < \theta \\ 1, & t \geq \theta \end{cases}.$$

Таку функцію називають виродженим розподілом, що надзвичайно рідко зустрічається на практиці. Найбільш часто в розрахунках використовується пуассоновський потік вимог, для якого імовірність надходження k вимог за час t визначається за формулою

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

де λ – інтенсивність потоку (середня кількість вимог, що надходять за одиницю часу).

Проміжний тип вхідного потоку, з якого можна одержати перші два типи потоків, описується ерлангенською щільністю розподілу моментів

часу між двома надходженнями вимог:

$$f(t) = \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t},$$

де k – параметр розподілу.

Цей розподіл перетворюється у вироджений при $k \rightarrow \infty$ і збігається з пуассоновським при $k = 0$.

Ерлангенський потік легко отримується з пуассоновського шляхом його розрідження (просіювання), коли пропускається не кожна вимога пуассоновського потоку, а тільки кожна k -а (рис.1.6).

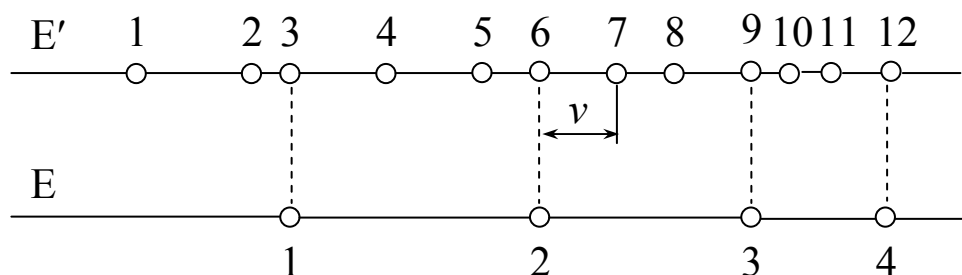


Рисунок 1.6 – Реалізація ерлангенського потоку

Вхідний потік може описуватися в деяких випадках нормальним законом розподілу. Відзначимо, що якщо середнє значення λ в пуассоновському потоці є великим, то такий потік можна описати законом, близьким до нормального, щільність розподілу якого визначена виразом

$$f(t) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(t-m_t)^2}{2\sigma_t^2}}.$$

де $m_t = \frac{1}{\lambda}$, а σ_t^2 – дисперсія довжини проміжків часу між появами сусідніх вимог.

Можна показати, що нормальний закон розподілу може бути хорошим наближенням до ерлангенського, якщо в ньому $k > 5$.

Потоки вимог можуть також описуватися й іншими законами розподілу (гіперекспоненціальним, Вейбулла, рівномірної щільності тощо).

Встановити закон розподілу можна шляхом аналізу великого статистичного матеріалу. При цьому будують гістограми й оцінюють закон за допомогою різних критеріїв (наприклад, Пірсона, Колмогорова тощо).

Якщо відсутній статистичний матеріал, то гіпотезу про закон розподілу потоку вимог можна висунути з фізичного аналізу розглянутого процесу. Наприклад, відомо, що нормальний закон розподілу проявляється в

основному тоді, коли на процес впливає безліч різних за змістом факторів, але кожен з яких вносить відносно малу частку в остаточний результат, тобто відсутні домінуючі фактори. Бета-розподіл крім наявності великої кількості випадкових факторів, кожен з яких чинить окремий вплив, характеризується також наявністю декількох факторів, вплив яких є істотним. Пуассоновський розподіл виявляється тоді, коли з великої кількості випадкових факторів істотно впливає на процес лише один з них. Не применшуючи значення потоків, що описуються іншими законами розподілу, зупинимося на розгляді найпростішого потоку, оскільки він широко застосовується при різних дослідженнях і розрахунках.

1.4.2 Найпростіший потік і його властивості

Найпростішим потоком називається потік, який характеризується трьома основними властивостями: стаціонарністю, відсутністю післядії та ординарністю.

Випадковий потік називається стаціонарним, якщо імовірність $P_n(\tau)$ надходження n вимог протягом часового відрізка τ не залежить від розташування цього відрізка на часовій осі, а залежить лише від його величини та інтенсивності потоку λ . При цьому відрізки не повинні накладатися один на одного, тобто якщо на часовій осі відкласти однакові, але такі, що не перетинаються, відрізки часу τ (рис. 1.7), то імовірність появи в цих інтервалах певної кількості вимог залежить від величини τ і не залежить від розміщення інтервалу на часовій осі (від моментів часу $t_1, t_2, t_3, \dots$).

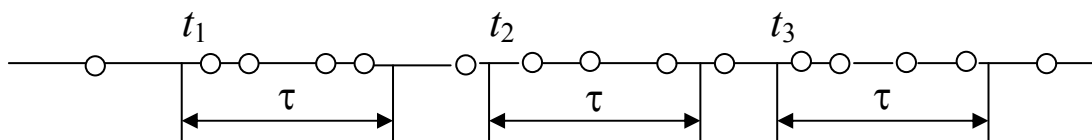


Рисунок 1.7 – Реалізація випадкового стаціонарного потоку

Два найпростіших однорідних потоки відрізняються один від одного тільки параметрами. Тому, щоб задати найпростіший потік, достатньо задати тільки його параметр λ .

Ординарність потоку вимог означає, що імовірність появи в один і той самий момент часу двох і більше вимог є нескінченно малою величиною, тобто $\frac{P_{>1}(\Delta t)}{\Delta t} \rightarrow 0$ при $\Delta t \rightarrow 0$.

Відсутність післядії полягає в тому, що імовірність надходження за відрізок τ певної кількості вимог n не залежить від того, скільки вимог уже надійшло в систему, тобто не залежить від передісторії досліджуваного явища.

Найпростіший потік знаходить широке застосування при дослідженнях різних великих систем і процесів. Це пояснюється таким:

- велика кількість робіт з дослідження операцій показує, що багато вхідних потоків добре узгоджуються з найпростішим потоком (потоки телефонних викликів, потоки потягів, що прибувають на залізничні вузли, морських суден, що прибувають у порт тощо);

- для інших типів потоків поки що не отримані прості аналітичні вирази для кількісного оцінювання якості функціонування систем;

- найпростіший потік є для систем найбільш складним для обслуговування. Тому, якщо засоби обслуговування розраховані на найпростіший потік, то обслуговування інших видів потоків при тій самій частоті надходження вимог буде більш якісним (час перебування в системі і черга – менші);

- при додаванні декількох ($N > 10$) випадкових потоків утворюється сумарний потік, що за своїми властивостями наближається до найпростішого, тобто якщо від великої кількості незалежних джерел надходять вимоги в одну систему, то такий потік можна розглядати як найпростіший незалежно від того, яким законом розподілу описується кожен з цих потоків.

Розглянемо математичну модель найпростішого потоку.

Нехай події E (надходження вимог) надходять одна за одною через випадкові проміжки часу (див. рис. 1.5, б). Кількість надходжень події E протягом інтервалу часу τ є випадковою величиною $N(\tau)$. Імовірність того, що $N(\tau)=n$, позначимо $P_n(\tau)$ і знайдемо її за допомогою властивостей найпростішого потоку.

Для будь-якого проміжку τ можна записати:

$$P_0(\tau) + P_1(\tau) + P_2(\tau) + \dots + P_K(\tau) + \dots = 1.$$

Цей же вираз справедливий для проміжку часу $\Delta\tau \rightarrow 0$:

$$P_0(\Delta\tau) + P_1(\Delta\tau) + P_2(\Delta\tau) + \dots + P_K(\Delta\tau) + \dots = 1.$$

Відповідно до властивості ординарності найпростішого потоку,

$$P_0(d\tau) + P_1(d\tau) = 1. \quad (1.2)$$

Імовірність того, що подія E відбудеться один раз у нескінченно малому інтервалі $d\tau$, відповідно до властивості стаціонарності буде дорівнювати:

$$P_1(d\tau) = \lambda d\tau.$$

З (1.2) випливає:

$$P_0(d\tau) = 1 - P_1(d\tau) = 1 - \lambda d\tau.$$

Визначимо тепер імовірність $P_n(\tau + d\tau)$. Для того, щоб в інтервалі $\tau + d\tau$ подія E настала n раз, повинен здійснитися один з двох несумісних випадків.

Перший випадок: n подій E з'явилося в інтервалі τ і 0 подій за час $d\tau$, що слідує безпосередньо за інтервалом τ .

Імовірність появи цього випадку дорівнює $P_n(\tau)(1 - \lambda d\tau)$.

Другий випадок: $n - 1$ подій E з'явилося в інтервалі τ і одна подія E в інтервалі $d\tau$, що слідує за τ . Відповідна імовірність дорівнює $P_{n-1}(\tau)\lambda d\tau$. На підставі теореми додавання можна записати:

$$P_n(\tau + d\tau) = P_n(\tau)(1 - \lambda d\tau) + P_{n-1}(\tau)\lambda d\tau + 0(d\tau), \quad (1.3)$$

де $0(d\tau)$ – у порівнянні з $d\tau$ нескінченно мала величина, яку можна не враховувати.

Вираз (1.3) можна переписати у вигляді:

$$P_n(\tau + d\tau) - P_n(\tau) = -\lambda P_n(\tau)d\tau + \lambda P_{n-1}(\tau)d\tau. \quad (1.4)$$

Розділивши ліву і праву частини (1.4) на $d\tau$, одержимо похідну:

$$P'_n(\tau) = -\lambda[P_n(\tau) - P_{n-1}(\tau)], \quad n > 0. \quad (1.5)$$

Знайдемо відповідну імовірність для $n=0$.

Для того, щоб подія E не з'явилася в інтервалі $\tau + d\tau$ жодного разу, достатньо, щоб вона не з'явилася ані в інтервалі τ , ані в інтервалі $d\tau$. Імовірність цієї події дорівнює $P_0(\tau)(1 - \lambda d\tau)$. Отже, можна записати:

$$P_0(\tau + d\tau) = P_0(\tau)(1 - \lambda d\tau) + 0(d\tau). \quad (1.6)$$

Перетворивши (1.6) аналогічно до (1.4), одержимо:

$$P'_0(\tau) = -\lambda P_0(\tau). \quad (1.7)$$

Об'єднавши (1.5) і (1.7), одержимо систему рівнянь:

$$\begin{cases} P'_0(\tau) = -\lambda P_0(\tau), \\ P'_n(\tau) = -\lambda[P_n(\tau) - P_{n-1}(\tau)], \quad n > 0. \end{cases} \quad (1.8)$$

Система лінійних диференціальних рівнянь (1.8) є системою станів для описуваного процесу. Це математична модель станів найпростішого потоку.

Доповнимо (1.8) очевидними початковими умовами:

$$P_0(0) = 1, \quad (1.9)$$

$$P_n(0) = 0, \quad n > 0. \quad (1.10)$$

Розв'язком першого рівняння (1.8) за умов (1.9), (1.10) є $P_0(\tau) = e^{-\lambda\tau}$.

Для знаходження розв'язку системи рівнянь (1.8) зручно використати перетворення Карсона-Лапласа: $f(S) = S \int_{-\infty}^{\infty} e^{-St} f(t) dt$, де S – символ перетворення Лапласа (комплексна змінна).

Встановимо $P_n^*(S) = LP_n(\tau)$, $n = 0, 1, 2, \dots$, де L – оператор перетворення Карсона-Лапласа.

За умов (1.9) і (1.10) рівняння системи (1.8) приводяться до вигляду:

$$P_0^*(S) = \frac{S}{\lambda + S}, \quad P_n^*(S) = \frac{\lambda}{\lambda + S} P_{n-1}^*(S), \quad (1.11)$$

що можна подати у вигляді рекурентних формул:

$$P_0^*(S) = \frac{S}{\lambda + S}; \quad P_1^*(S) = \frac{\lambda S}{(\lambda + S)^2}; \quad P_2^*(S) = \frac{\lambda^2 S}{(\lambda + S)^3}; \quad \dots \quad P_n^*(S) = \frac{\lambda^n S}{(\lambda + S)^{n+1}}.$$

За допомогою таблиць перетворення Лапласа (чи методом відрахувань) знаходимо оригінал (1.11):

$$P_n(\tau) = \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau}, \quad n = 0, 1, 2, \dots \quad (1.12)$$

Дане співвідношення являє собою формулу Пуассона, тобто найпростіший потік описується пуассонівським розподілом, припущення про ординарність якого запишуться у вигляді:

$$P_0(d\tau) = 1 - \lambda d\tau;$$

$$P_1(d\tau) = \lambda d\tau.$$

У кінцевому результаті найпростіший потік може бути визначений як точковий марковський процес з матрицею переходу:

	E_0	E_1	E_2	E_3	стан в момент $t + d\tau$
E_0	$1 - \lambda d\tau$	$\lambda d\tau$	0	0	...
E_1	0	$1 - \lambda d\tau$	$\lambda d\tau$	0	...
$F = E_2$	0	0	$1 - \lambda d\tau$	$\lambda d\tau$	...
E_3	0	0	0	$1 - \lambda d\tau$	...
стан в момент t	0	0	0	0	...

Граф переходів, що відповідає процесу Пуассона, відображає суть процесу: для усіх n можливий перехід зі стану E_n у E_{n+1} з імовірністю $\lambda d\tau$ або цей стан E_n зберігається з імовірністю $1 - \lambda d\tau$ (рис. 1.8).

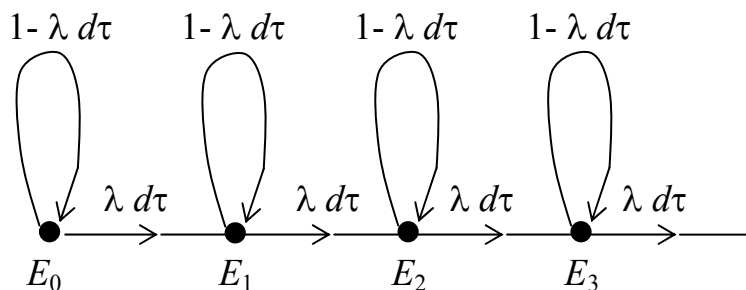


Рисунок 1.8 – Граф переходів точкового марковського процесу

Будь-який потік вимог утворить імовірнісний процес, який у всіх випадках являє собою послідовність подій, що настають через випадкові інтервали при неперервному відліку часу.

Кількісною характеристикою потоку є його інтенсивність λ , що визначається як математичне сподівання кількості вимог, що надходять за одиницю часу.

В реальних умовах для великих відрізків часу рідко виконується умова стаціонарності потоку й інтенсивність потоку змінюється, в одних випадках, за деяким неперіодичним законом, а в інших, коли досліджуються процеси сезонного характеру – за періодичним законом.

Однак за певних умов можна припустити прояву пуассонівського процесу протягом деяких інтервалів часу роботи СМО, вважаючи, що інтенсивність потоку λ залишається постійною на певних відрізках часу $t_1, t_2, t_3, t_4 \dots$ (рис. 1.9) Це дає можливість звести вхідний потік до потоку Пуассона на окремих часових ділянках зі своїм значенням λ , що суттєво спрощує аналіз роботи системи.

Визначимо математичне сподівання кількості вимог найпростішого потоку за проміжок $(0, t)$:

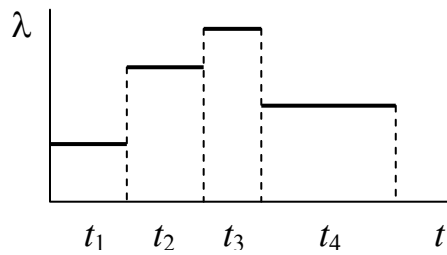


Рисунок 1.9 – Подання пуассонівського процесу

$$M_t[k] = \sum_{k=0}^{\infty} k P_k(t) = e^{-\lambda t} \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} = e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} = \lambda t.$$

Отже, інтенсивність потоків згідно визначення буде $M_1[k] = \lambda$.

Параметр λ ще називають щільністю потоку вимог. Одна з властивостей найпростішого потоку – його ординарність. Однак на практиці нерідко трапляються випадки, коли на обслуговування надходить одночасно група вимог, наприклад, надходження техніки в ремонт. Це так зване групове надходження вимог. Для розв’язання подібних задач у ТМО розроблено спеціальний математичний апарат, який у даному посібнику не розглядається.

1.4.3 Розподіл інтервалів часу між двома подіями найпростішого потоку

Розглянемо найпростіший потік (процес Пуассона) і визначимо закон розподілу інтервалів часу між двома послідовними подіями (надходженнями вимог). Як відомо, закон розподілу (функція розподілу) часу між двома сусідніми подіями є найважливішою характеристикою будь-якого потоку.

Нехай θ – випадкова величина, що характеризує довжину цих інтервалів, $f(v)$ – щільність розподілу, а $F(v)$ – функція надійності, що характеризує імовірність того, що $\theta > v$. Однак $F(v)$ являє собою умовну імовірність того, що подія E не відбудеться в інтервалі $(t, t+v)$, якщо вона відбулася в момент t . Ця умовна імовірність дорівнює відповідній безумовній імовірності того, що за час v не відбудеться подія E , що можна знайти з (1.12), встановивши $k = 0$:

$$P_0(v) = \frac{(\lambda v)^0 e^{-\lambda v}}{0!} = e^{-\lambda v}.$$

Дана формула описує функцію надійності:

$$F(v) = P(\theta > v) = e^{-\lambda v}.$$

Функція розподілу пов'язана з функцією надійності очевидним співвідношенням:

$$F(t) = P(\theta \leq t) = 1 - p(\theta > v). \quad (1.13)$$

Продиференціювавши (1.13), отримаємо:

$$f(t) = \lambda e^{-\lambda t}. \quad (1.14)$$

Відзначимо, що внаслідок марковського характеру властивості розглянутого процесу інтервали θ між подіями E взаємно незалежні. З (1.14) видно, що розподіл інтервалів часу між двома подіями пуассонівського процесу описується експоненціальним законом, для якого визначимо математичне сподівання і дисперсію:

$$\bar{\theta} = M[\theta] = \int_0^{\infty} v f(v) dv = \frac{1}{\lambda}. \quad (1.15)$$

Величина $\frac{1}{\lambda}$ визначає собою середнє значення довжини інтервалу між двома послідовними подіями.

Для визначення дисперсії довжини інтервалів θ скористаємося формулою гамма-функції $\Gamma(n) = (n-1)! = \int_0^{\infty} x^{n-1} e^{-x} dx$. Тоді

$$\sigma_v^2 = \int_0^{\infty} v^2 f(v) dv - [M(\theta)]^2 = \frac{1}{\lambda} \int_0^{\infty} (\lambda t)^2 e^{-\lambda t} d\lambda t - \frac{1}{\lambda^2} = \frac{\Gamma(3)}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Можна легко довести, що точковий процес, в якому інтервали між подіями взаємно незалежні і належать одному і тому ж показниковому закону (1.14), є пуассонівським процесом.

Твердження про те, що процес є пуассонівським за розподілом кількості подій (1.12) чи що розподіл тривалості інтервалів процесу показниковий – рівносильні.

Наведений висновок справедливий не тільки стосовно вхідних потоків вимог, але і процесу обслуговування.

1.4.4 Основні характеристики обслуговуючих приладів

Основна характеристика обслуговуючих приладів – їх продуктивність, за якою можна судити про пропускну здатність різних систем. Однак для дослідження систем масового обслуговування особливий інтерес викликає час обслуговування вимог, що надходять до приладу, оскільки цей час, в свою чергу, визначається продуктивністю приладу.

Оскільки час обслуговування вимог є майже завжди випадковою величиною, для оцінювання пропускну здатності СМО необхідно знати його закони розподілу. Тому повною характеристикою часу обслуговування є його функція розподілу чи щільність розподілу.

У дослідженнях СМО використовуються експоненціальний, ерлангенський, а також нормальний та регулярний закони розподілу часу обслуговування, про які вже згадувалося. При розв'язанні практичних і теоретичних задач найчастіше застосовується експоненціальний (показниковий) розподіл, що істотно зменшує складність аналізу СМО і дозволяє отримати зручні для інженерної практики розрахункові вирази. Розв'язання задач ТМО з довільним законом розподілу часу обслуговування реалізується з великими труднощами.

При показниковому законі функція розподілу має вигляд

$$F(t) = 1 - e^{-\mu t},$$

де $\mu = \frac{1}{\bar{t}_{об}}$ – позитивна постійна величина, що називається інтенсивністю

обслуговування;

$\bar{t}_{об}$ – математичне сподівання часу обслуговування.

Щільність розподілу тривалості обслуговування $f(t) = \mu e^{-\mu t}$.

Для показникового розподілу середнє значення і стандартне відхилення часу обслуговування дорівнює $1/\mu$. Якщо випадкова величина $t_{об}$ є тривалість обслуговування, то додаткова функція розподілу (функція надійності)

$$F^*(v) = P(t_{об} > v) = \int_{t_0}^{\infty} \mu e^{-\mu v} dv = e^{-\mu t_0}.$$

За допомогою цього виразу можна обчислити імовірність того, що обслуговування вимоги буде закінчено протягом інтервалу часу Δt за умови, що обслуговування триває протягом часу t_0 . Ця умовна імовірність дорівнює

$$P\{t_0 \leq t_{об} \leq t_0 + \Delta t / t_{об} \geq t_0\} = \frac{P\{t_0 \leq t_{об} \leq t_0 + \Delta t\}}{P\{t_{об} \geq t_0\}} \approx \frac{\mu e^{-\mu t} \Delta t}{e^{-\mu t}} = \mu \Delta t. (1.16)$$

З (1.16) видно, що імовірність завершення обслуговування протягом малого інтервалу часу постійна і не залежить від того, як довго тривало обслуговування. Таким чином, показникове обслуговування може трактуватися як зовсім випадкова операція.

Показниковий закон розподілу часу обслуговування припускає, що значна частина вимог буде обслуговуватися швидко за дуже короткий інтервал часу. Це не завжди виконується на практиці. З суто інтуїтивних міркувань випливає, що більшість операцій з обслуговування різних вимог не можуть бути виконані за час, менший деякого значення t_{min} . Тоді показниковий розподіл часу обслуговування стає малопридатним, а тому для опису тривалості обслуговування вимог застосовують так званий двопараметричний експоненціальний розподіл, щільність якого описується виразом

$$f(t) = \begin{cases} \mu e^{-\mu(t-t_{min})} & , t > t_{min} \\ 0 & , t \leq t_{min} \end{cases}$$

де μ і t_{min} – параметри закону розподілу.

Цей вираз відрізняється від звичайного показникового розподілу наявністю другого параметра t_{min} .

Щільність розподілу часу обслуговування можна визначити за формулою Ерланга:

$$f_k(t) = \frac{(\mu k)^k}{\Gamma(k)} e^{-\mu k t} t^{k-1} \quad \text{при } t > 0,$$

$$f_k(t) = 0 \quad \text{при } t < 0,$$

де $\Gamma(k) = (k-1)!$ – гамма-функція.

Можна показати, що $f_k(t)$ являє собою щільність розподілу суми k незалежних випадкових величин з показниковим законом розподілу.

СМО при показниковому законі розподілу має одну важливу властивість, яку потрібно враховувати при оцінюванні ефективності її роботи.

Нехай у СМО, що складається з S різнотипних приладів, надійшла вимога, яку починають обслуговувати всі канали одночасно.

Час обслуговування кожним каналом підпорядкований показнико-

вому закону розподілу з параметром μ_i ($i = \overline{1, S}$). Обслуговування закінчується, як тільки один із приладів виконав свою задачу. Для цього випадку закон розподілу буде також показниковим і описується виразом:

$$F(t_{об} < t) = 1 - e^{-(\mu_1 + \mu_2 + \dots + \mu_S)t}$$

з параметром $\mu = \sum_{i=1}^S \mu_i$.

Таким чином, для характеристики каналів необхідно знати закон розподілу часу обслуговування вимог і інтенсивність обслуговування каналу μ , яка являє собою середню кількість вимог, що обслуговуються за обрану одиницю часу. При цьому варто звернути увагу на те, що обрана одиниця часу для визначення μ повинна бути такою ж, як і для визначення інтенсивності потоку λ .

Питання для самоконтролю

1. Дайте означення теорії масового обслуговування і покажіть на прикладах виробничих систем її зв'язок з матеріалом випадкових процесів.
2. Дайте означення системи масового обслуговування і наведіть приклади з повсякденного життя.
3. Перелічіть основні структури систем масового обслуговування і поясніть їх особливості.
4. Запишіть основні показники ефективності систем масового обслуговування; проведіть аналогію між ними та числовими характеристиками випадкових величин.
5. Перелічіть основні закони розподілу потоків вимог; запишіть їх вирази; покажіть взаємозв'язок цих виразів.
6. Перелічіть і поясніть основні властивості найпростішого потоку; виведіть самостійно остаточний вираз, що описує цей потік.
7. Поясніть причини широкого застосування на практиці найпростішого потоку.
8. Назвіть закон розподілу інтервалів часу між моментами появи вимог пуассонівського потоку.
9. Перелічіть і запишіть основні закони розподілу часу обслуговування. Назвіть основні причини широкого застосування показникового розподілу.
10. Запишіть скорочене позначення системи, що складається з трьох приладів, на яку надходить ерлангенський потік вимог, час обслуговування яких описується регулярним і довільним законами розподілу. У системі допускається черга не більш як з 20 вимог.

2 МАРКОВСЬКІ МОДЕЛІ СИСТЕМ МАСОВОГО

ОБСЛУГОВУВАННЯ

Вивчення систем масового обслуговування слід почати з марковських систем, для яких характерна відсутність післядії, що спрощує одержання розрахункових залежностей, необхідних в інженерній практиці, хоча і вимагає для ряду марковських систем трудомістких обчислень. Найбільш простою є модель одноканальної СМО, на прикладі якої досить докладно повинні бути розглянуті методи аналізу марковських систем, що дозволить надалі обмежитися посиланням на ці методи. При розгляді моделей завжди потрібно звертати увагу на умови чи припущення, на основі яких будується модель, тому що при їх зміні модель вийде іншою.

2.1 Система $M/M/1/\infty$

Графічна модель системи $M/M/1/\infty$ отримується з моделі, показаної на рис.1.1, при $S=1$. Така СМО аналізується при таких припущеннях (умовах).

Система містить у собі один обслуговуючий канал, на який надходить найпростіший потік вимог з параметром λ , час обслуговування вимог розподілено за показниковим законом з параметром μ ; вимоги обслуговуються в порядку надходження до системи (першим прийшов – першим обслуговується), довжина черги не обмежується.

Аналіз системи проводиться також при усталеному (стаціонарному) режимі, що забезпечується умовою

$$\psi = \frac{\lambda}{\mu} = \lambda \overline{t_{об}} < 1,$$

де ψ – завантаження системи, що являє собою основну характеристику процесу масового обслуговування.

Слід одразу звернути увагу на те, що при значенні $\psi \geq 1$ у системі утвориться черга нескінченної довжини. Це означає неможливість такої системи обслуговувати вимоги, що надходять.

Для аналізу марковської системи використовують два методи. Перший метод полягає в тому, що за допомогою складених диференціальних рівнянь знаходять усі числові характеристики якості роботи системи. Цей метод називають диференціальним. Другий метод полягає в знаходженні числових характеристик системи за допомогою складених інтегральних рівнянь. Тому метод називають інтегральним.

2.1.1 Диференціальний метод аналізу (рівняння станів)

Основне завдання методу полягає у визначенні розподілу імовірностей випадкової величини n , яка характеризує кількість вимог у системі при таких припущеннях, що впливають з властивостей найпростішого потоку:

а) імовірність того, що вимога надходить у систему в інтервалі Δt , є нескінченно мала величина порядку Δt : вона дорівнює $\lambda \Delta t$;

б) імовірність того, що обслуговування однієї вимоги закінчиться в інтервалі Δt , є нескінченно мала величина порядку Δt : вона дорівнює $\mu \Delta t$;

в) імовірність надходження декількох вимог чи декількох закінчень обслуговування в інтервалі Δt є нескінченно мала величина $0(\Delta t)$, яка не враховується.

На основі даних припущень легко скласти диференціальні рівняння, що описують процес обслуговування, за допомогою яких знаходять вирази для визначення імовірнісних числових характеристик якості обслуговування. Для цього розглядається імовірність $P_n(t+\Delta t)$ того, що в системі знаходиться n вимог ($n > 0$) у момент $(t+\Delta t)$. Ця імовірність дорівнює сумі імовірностей чотирьох несумісних подій, сутність яких впливає з їх описання.

Подія 1. Імовірність появи цієї події визначається добутком імовірностей таких ситуацій:

а) у момент t в системі було n вимог, імовірність чого є $P_n(t)$;

б) в інтервалі Δt не було надходження вимог, імовірність чого є $(1-\lambda \Delta t)$;

в) в інтервалі Δt жодне обслуговування не закінчилося, імовірність чого є $(1-\mu \Delta t)$.

Подія 2. Імовірність появи цієї події визначається добутком імовірностей таких ситуацій:

а) в момент t у системі знаходилося $(n+1)$ вимог, імовірність чого – $P_{n+1}(t)$;

б) в інтервалі Δt не було жодних надходжень вимог, імовірність чого – $(1-\lambda \Delta t)$;

в) в інтервалі Δt закінчилося одне обслуговування, імовірність чого – $\mu \Delta t$.

Подія 3. Імовірність появи цієї події визначається добутком імовірностей таких подій:

а) в момент t у системі знаходилося $(n-1)$ вимог – $P_{n-1}(t)$;

б) в інтервалі Δt надійшла одна вимога – $\lambda \Delta t$;

в) в інтервалі Δt жодне обслуговування не закінчилося – $(1-\mu \Delta t)$.

Подія 4. Імовірність появи цієї події визначається добутком імовірностей таких подій:

а) в момент t у системі було n вимог – $P_n(t)$;

б) в інтервалі Δt надійшла одна вимога – $\lambda \Delta t$;

в) в інтервалі Δt закінчилося одне обслуговування – $\mu \Delta t$.

Вказані імовірності появи подій можуть бути записані в такий спосіб:

$$\begin{aligned} P_n(t) (1-\lambda\Delta t) (1-\mu\Delta t) &= P_n(t) (1-\lambda\Delta t-\mu\Delta t)+O_1(\Delta t); \\ P_{n+1}(t) (1-\lambda\Delta t) (\mu\Delta t) &= P_{n+1}(t) \mu\Delta t+O_2(\Delta t); \\ P_{n-1}(t) (\lambda\Delta t) (1-\mu\Delta t) &= P_{n-1}(t) \lambda\Delta t+O_3(\Delta t); \\ P_n(t) (\lambda\Delta t) (\mu\Delta t) &= O_4(\Delta t). \end{aligned} \quad (2.1)$$

Величини $O_i(\Delta t)$ ($i = 1, 2, 3, 4$) є нескінченно малими, але більш високих порядків у порівнянні з Δt при $t \rightarrow 0$.

Підсумувавши імовірності (2.1), а потім перенісши в ліву частину рівняння суми імовірностей $P_n(t)$ і розділивши обидві частини рівняння на Δt при $\Delta t \rightarrow 0$, одержимо:

$$\frac{d}{dt} P_n(t) = \lambda P_{n-1}(t) + \mu P_{n+1}(t) - (\lambda + \mu) P_n(t), \quad n > 0. \quad (2.2)$$

До рівняння (2.2) необхідно додати рівняння, що відповідає випадку, коли в системі в момент $t + \Delta t$ немає жодної вимоги.

Імовірність того, що в системі в момент часу $t + \Delta t$ вимог немає, дорівнює сумі імовірностей двох таких несумісних подій.

1. Добутки імовірностей подій:

а) в момент t у системі не знаходилося жодної вимоги, імовірність чого є $P_0(t)$;

б) в інтервалі Δt не надійшло жодної вимоги, імовірність чого є $(1-\lambda\Delta t)$.

2. Добутки імовірностей подій:

а) в момент t у системі знаходилася одна вимога – $P_1(t)$;

б) в інтервалі Δt не надійшло жодної вимоги – $(1-\lambda\Delta t)$;

в) в інтервалі Δt закінчилося одне обслуговування – $\mu\Delta t$.

Виконавши перетворення аналогічно перетворенням (2.1), остаточно одержимо:

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t) + \mu P_1(t). \quad (2.3)$$

Рівняння (2.2) і (2.3) являють собою математичну модель одноканальної СМО при найпростішому потоці й експоненціальній тривалості обслуговування, записану в диференціальній формі. За допомогою цієї моделі легко знайти всі числові характеристики для сталого режиму роботи СМО.

2.1.2 Визначення імовірнісних і моментних характеристик СМО

при сталому режимі

Для системи існує сталий режим (або вона статистично стійка), якщо при $\Delta t \rightarrow \infty$ для кожного n існує границя $\lim_{t \rightarrow \infty} P_n(t) = P_n$, ($n=0, 1, 2, \dots$), причому принаймні одне із значень P_n не дорівнює нулю.

При $t \rightarrow \infty$ похідні, що входять у (2.2) і (2.3), наближаються до нуля, і рівняння приводяться до вигляду:

$$\lambda p_0 = \mu p_1; \quad (2.4)$$

$$(\lambda + \mu)p_n = \lambda p_{n-1} + \mu p_{n+1}. \quad (2.5)$$

На основі (2.4) і (2.5) можна записати:

$$P_0 = p_0; P_1 = \psi p_0; P_2 = \psi^2 p_0; \dots P_n = \psi^n p_0. \quad (2.6)$$

Додавши ліві і праві частини (2.6), отримаємо:

$$\sum_{n=0}^{\infty} p_n = p_0 \sum_{n=0}^{\infty} \psi^n. \quad (2.7)$$

Оскільки

$$\sum_{n=0}^{\infty} \psi^n = \frac{1}{1-\psi}, \quad \sum_{n=0}^{\infty} p_n = 1,$$

то (2.7) приводиться до рівності: $1 = p_0 \frac{1}{1-\psi}$, звідки одержуємо

$$P_0 = 1 - \psi. \quad (2.8)$$

Підставляючи (2.8) у (2.6), визначаємо

$$P_n = \psi^n (1 - \psi). \quad (2.9)$$

За допомогою (2.9) можна легко знайти функцію розподілу імовірностей того, що ($N \leq n$):

$$\begin{aligned} \sum_{a=0}^n p_a &= \sum_{a=0}^n \psi^a (1 - \psi) = (1 - \psi) \sum_{a=0}^n \psi^a = \\ &= (1 - \psi)(1 + \psi + \psi^2 + \dots + \psi^n) = (1 - \psi) \frac{1 - \psi^{n+1}}{1 - \psi} = 1 - \psi^{n+1}. \end{aligned}$$

Отже, $P(N \leq n) = 1 - \psi^{n+1}$.

Імовірність того, що кількість вимог у системі перевищить n , дорівнює

$$P(N > n) = 1 - (1 - \psi)^{n+1} = \psi^{n+1}. \quad (2.10)$$

З (2.10) видно, що імовірність перебування в системі щонайменше однієї вимоги визначається величиною завантаження $P(N > 0) = \psi$.

Отримані вирази імовірнісних характеристик дозволяють визначити формули для розрахунку моментних характеристик СМО. Для цього необхідно в загальні вирази моментних характеристик, розглянуті раніше, підставити імовірність p і виконати звичайні перетворення. В результаті знаходимо:

- середня кількість вимог у системі:

$$\begin{aligned} \bar{n} = M[N] &= \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n (1 - \psi) \psi^n = \\ &= (1 - \psi) [\psi + 2\psi^2 + 3\psi^3 + \dots] = \frac{\psi}{1 - \psi}; \end{aligned} \quad (2.11)$$

- середня кількість вимог у черзі:

$$\bar{v} = \sum_{n=2}^{\infty} (n - 1) p_n = \frac{\psi^2}{1 - \psi}. \quad (2.12)$$

Середню тривалість перебування вимог у системі й у черзі можна одержати з огляду на таке: у середньому за одиницю часу через систему проходить $\lambda = (1 - p_0)\mu$ вимог. Це випливає з рівності $\psi = \lambda/\mu$ і рівняння (2.8) або таких міркувань: середня кількість вимог, що надходять за час T , дорівнює λT ; середня кількість вимог, що залишають систему, дорівнює добутку μ на ту частину інтервалу T , протягом якої відбувається обслуговування і яка дорівнює $(1 - p_0)T$. Звідси

$$\lambda = (1 - p_0)\mu. \quad (2.13)$$

Рівняння (2.13) називається рівнянням витрати і дозволяє швидко знайти величину p_0 .

Для середнього часу перебування вимог у системі, з урахуванням цих міркувань, можна записати очевидний вираз:

$$\bar{u} = \frac{\bar{n}}{(1 - p_0)\mu} = \frac{\bar{n}}{\lambda} = \frac{1}{\lambda} \frac{\psi}{1 - \psi} = \frac{1}{\mu} \frac{1}{1 - \psi}. \quad (2.14)$$

Аналогічно для середнього часу очікування в черзі отримаємо:

$$\bar{\omega} = \frac{\bar{v}}{(1-p_0)\mu} = \frac{\bar{v}}{\lambda} = \frac{1}{\mu} \frac{\psi}{1-\psi} . \quad (2.15)$$

Вирази (2.14) і (2.15) пов'язані між собою очевидним співвідношенням:

$$\bar{u} = \bar{\omega} + \frac{1}{\mu} . \quad (2.16)$$

З наведених виразів (2.11)-(2.16) видно, що для визначення всіх основних характеристик якості функціонування системи достатньо визначити лише одну, за допомогою якої при очевидних взаємозв'язках між характеристиками можна легко отримати інші.

Використовуючи вирази (2.11)-(2.16), можна побудувати графіки залежностей значень характеристик від величини завантаження. Один з таких графіків показаний на рис.2.1.

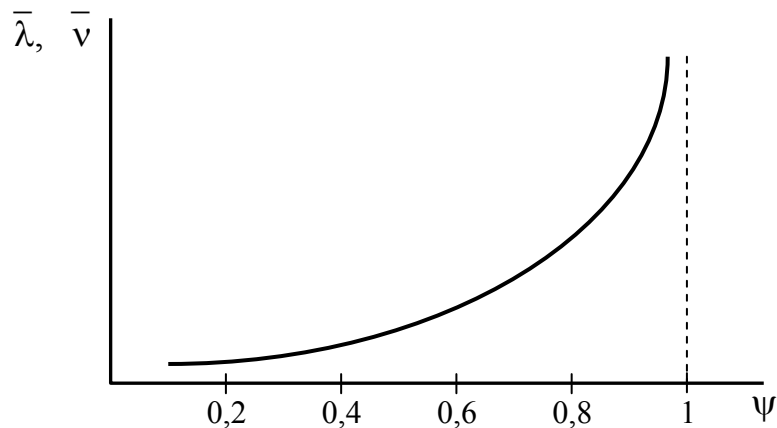


Рисунок 2.1 - Графік залежності значень характеристик від величини завантаження

2.1.3 Інтегральний метод аналізу СМО

Метод дозволяє знайти вираз для функцій розподілу часу перебування вимог у системі $G(x)$ і в черзі $F(x)$, за допомогою яких визначаються моментні характеристики СМО. Розглянемо його сутність і необхідні залежності, що будуть потрібні надалі [1-3].

Сутність інтегрального методу зводиться до того, що розглядаються моменти надходження кожної n -ї вимоги в систему, і моменти її виходу із системи ($n=1, 2, \dots$). Для кожної вимоги визначаються $F(x)$ і $G(x)$. Потім розглядається сталий режим роботи, для якого використовують такі інтегральні рівняння:

$$F(x) = \int_0^{\infty} a(\xi) G(x + \xi) d\xi, \quad (2.17)$$

$$G(x) = \int_0^x S(\xi) F(x - \xi) d\xi, \quad (2.18)$$

де $a(\xi) = \lambda e^{-\lambda \xi}$ – функція щільності інтервалів часу між моментами надходження вимог пуассонівського потоку з параметром λ в систему;

$S(\xi)$ – функція щільності часу обслуговування вимог з параметром μ .

Оскільки одержати розв’язання в явному вигляді за допомогою (2.17) і (2.18) неможливо, знаходять їх зображення, які для перетворень Карсона-Лапласа в умовах пуассонівського потоку мають відповідний вигляд:

$$F(p) = \frac{pF(0) - \lambda G(p)}{p - \lambda}, \quad (2.19)$$

$$G(p) = [1 - S'(p)] / F(p), \quad (2.20)$$

де p – комплексна змінна в перетвореннях Карсона-Лапласа;

$$S'(x) = \int_x^{\infty} S(\xi) d\xi.$$

Для марковських СМО $S(\xi) = \mu e^{-\mu \xi}$, а відповідне зображення

$$S'(p) = \frac{p}{p + \mu}. \quad (2.21)$$

Підставивши (2.20) у (2.19), одержимо:

$$F(p) = F(0) \frac{p}{p - \lambda S'(p)}. \quad (2.22)$$

Слід зазначити, що вираз (2.22) справедливий для довільного закону розподілу часу обслуговування вимог. При підставленні в нього (2.21) маємо:

$$F(p) = F(0) \frac{p + \mu}{p - \lambda + \mu}. \quad (2.23)$$

За допомогою таблиць перетворень Лапласа знаходимо оригінал:

$$F(x) = F(0) \left(\frac{1}{1-\psi} - \frac{\psi}{1-\psi} e^{-(\mu-\lambda)x} \right), \quad (2.24)$$

який при $x \rightarrow \infty$ буде дорівнювати одиниці. Підставивши в (2.24) вираз $F(0)=1-\psi$, одержимо остаточний результат:

$$F(x) = 1 - \psi e^{-(\mu-\lambda)x}. \quad (2.25)$$

За допомогою (2.25) знаходимо вираз для середнього часу очікування в черзі:

$$\bar{\omega} = \int_0^{\infty} x dF(x) = \frac{\psi}{\mu(1-\psi)}.$$

Підставивши (2.21) і (2.23) у (2.20), одержимо:

$$G(p) = \mu(1-\psi) \frac{1}{p - \lambda + \mu},$$

оригінал якого має вигляд:

$$G(x) = 1 - e^{-(\mu-\lambda)x}. \quad (2.26)$$

За допомогою (2.26) знаходимо уже відомий вираз для середнього часу перебування вимоги в системі:

$$u = \int_0^{\infty} x dG(x) = \frac{1}{\mu(1-\psi)}.$$

Таким чином, отримано всі необхідні розрахункові вирази для аналізу марковської одноканальної СМО. Однак для практики більш цікавим є закон розподілу вихідного зі СМО потоку вимог.

2.1.4 Вихідний потік вимог

Для аналізу багатофазних СМО і мереж масового обслуговування, а також для визначення параметра вхідного потоку і функції розподілу тривалості обслуговування вимог, необхідно знати закон розподілу вихідного зі СМО потоку, оскільки у зазначених системах він є вхідним для інших фаз обслуговування. Основна задача при цьому полягає у визначенні щільності розподілу $d(t)$ або функції розподілу інтервалів часу між моментами виходу вимог з фази (системи). Розглянемо СМО типу $M/G/1/\infty$, тобто систему, на яку надходить пуассонівський потік вимог з параметром λ . Ці вимоги обслуговуються за довільним законом розподілу $S(t)$.

Для визначення закону розподілу вихідного потоку зручно скористатися перетвореннями Лапласа. Позначимо $D(p)$ і $S(p)$ відповідні перетворення Лапласа щільності розподілу вихідного потоку і часу обслуговування вимог. Обчислимо $D(p)$ у момент, коли вимога залишає систему. При цьому можлива одна з двох подій: 1) у момент виходу вимоги із системи в черзі є хоча б одна вимога, що очікує обслуговування; 2) вимог, які очікують, немає. У першому випадку проміжок часу між вихідними зі СМО вимогами розподілений за тим самим законом, що і час обслуговування, тому справедливо $D(p)=S(p)$.

Якщо після виходу першої вимоги система виявилась порожньою, то інтервал між вимогою, що вийшла, і виходом наступної буде складатися з двох проміжків часу: перший проміжок – час до надходження в систему наступної (другої) вимоги і другий – час обслуговування цієї вимоги. Оскільки ці два проміжки розподілені незалежно, щільність розподілу імовірностей їх суми дорівнює згортці щільностей розподілу величин, що додаються. Перетворення Лапласа щільності розподілу суми дорівнює добутку вихідних щільностей розподілу. Для даного випадку

$$D(p) = A(p)S(p) = \frac{\lambda}{p + \lambda} S(p),$$

де $A(p)$ – перетворення Лапласа щільності розподілу інтервалів часу між вимогами вхідного потоку.

Оскільки імовірність застати систему зайнятою для вимоги, яка надходить, дорівнює ψ , а порожньою – $(1-\psi)$, то в загальному випадку можна записати:

$$D(p) = \psi S(p) + (1-\psi)A(p)S(p). \quad (2.27)$$

Для експоненціального закону розподілу часу обслуговування (2.27) приводиться до вигляду:

$$D(p) = \psi \left(\frac{\mu}{p + \mu} \right) + (1 - \psi) \left(\frac{\lambda}{p + \lambda} \right) \left(\frac{\mu}{p + \mu} \right),$$

і після елементарних перетворень зводиться до виразу:

$$D(p) = \frac{\lambda}{p + \lambda}. \quad (2.28)$$

З (2.28) випливає, що розподіл проміжків часу між вихідними вимогами $D(t) = 1 - e^{-\lambda t}$, $t \geq 0$.

Звідси робимо висновок: вихідний потік для системи $M / G / 1 / \infty$ є пуассонівським з параметром λ , тобто таким самим, як і вхідний потік.

Розглянемо тепер випадок довільного часу обслуговування $S(p)$. Для нього вираз (2.27) може бути приведений до вигляду:

$$D(p) = \frac{\lambda + p\psi}{p + \lambda} S(p) . \quad (2.29)$$

З (2.29) можна зробити висновок: для системи, на яку надходить пуассонівський потік вимог, вихідний потік описується законом, що залежить від закону розподілу часу обслуговування $S(p)$. Якщо цей закон не експоненціальний, то і вихідний потік не буде пуассонівським.

2.2 Багатоканальні системи масового обслуговування ($M/M/S/\infty$)

Розглянемо багатоканальну систему обслуговування, на яку надходить найпростіший потік вимог з параметром λ , а час обслуговування кожним каналом системи розподілено експоненціально з параметром μ . Структурна схема такої СМО показана на рис.1.1.

Припустимо, що має місце єдина черга і обслуговування вимог здійснюється в порядку черговості відразу ж після звільнення будь-якого каналу системи. Для дослідження багатоканальної системи будемо використовувати диференціальний метод, тому що при цьому процес обслуговування носить комбінаторний характер, для якого інтегральний метод незручний.

Будемо вважати, що в системі діють S незалежних один від одного каналів. Для багатоканальної системи назвемо сумарним завантаженням (або повним коефіцієнтом використання системи) співвідношення

$$\psi = \frac{\lambda}{\mu_0}, \text{ де } \mu_0 - \text{сумарна інтенсивність обслуговування, обумовлена кількі-}$$

стю каналів системи: $\mu_0 = S\mu$.

Як і раніше, будемо розглядати сталий режим, для якого $\psi < 1$.

Аналіз системи проведемо як для однакових, так і неоднакових каналів. Неоднаковими назвемо канали, які мають різну інтенсивність обслуговування одних і тих самих вимог, а однакові канали характеризуються інтенсивністю обслуговування μ , однаковою для всіх каналів.

2.2.1 Багатоканальна СМО з однаковими приладами

Аналіз багатоканальної СМО за допомогою диференціального методу проводиться аналогічно до одноканальної СМО. Тому відразу можна розглянути і навести без виведення рівняння станів системи для сталого режиму її роботи:

$$\begin{aligned} \lambda p_0 &= \mu p_1; \\ (\lambda + n\mu) p_n &= \lambda p_{n-1} + (n+1)\mu p_{n+1}, \quad 1 \leq n \leq S; \end{aligned}$$

$$(\lambda + S\mu) p_n = \lambda p_{n-1} + S\mu p_{n+1}, n > S.$$

З цих рівнянь знаходимо:

$$p_n = p_0 \frac{S^n \psi^n}{n!}, 1 \leq n < S;$$

$$p_n = p_0 \frac{S^S \psi^n}{S!}, n > S.$$

Імовірність p_0 знаходимо з умови $\sum_{n=0}^{\infty} p_n = 1$, яка може бути записана у вигляді:

$$p_0 \left[\sum_{n=0}^{S-1} \frac{S^n \psi^n}{n!} + \sum_{n=S}^{\infty} \frac{S^S \psi^n}{S!} \right] = 1. \quad (2.30)$$

З (2.30) знаходимо:

$$p_0 = \frac{1}{\frac{S^S \psi^S}{S!(1-\psi)} + \sum_{n=0}^{S-1} \frac{S^n \psi^n}{n!}}.$$

Можна показати, що $\lim_{S \rightarrow \infty} p_0 = e^{-\psi}$, тобто $p_0 \neq 1$ навіть при $S = \infty$. Це пояснюється тим, що n означає кількість вимог у черзі і (або) на обслуговуванні.

У розглянутих виразах завантаження системи ψ завжди менше одиниці. В той же час завантаження, розглянуте відносно одного приладу $\psi_{1n} = \lambda / \mu$, може значно перевищувати одиницю. Наведемо кілька значень p_0 для різних S і ψ_{1n} (див. табл. 2.1).

Таблиця 2.1

ψ_{1n}	$S=2$	$S=3$	$S=4$	$S=5$	$S=6$	$S=7$	...	$S=\infty$
1	0,333	0,363	0,367	0,367	0,367	0,367	...	0,36788
2		0,111	0,130	0,134	0,135	0,135	...	0,13534
3			0,037	0,046	0,049	0,049	...	0,04978
4				0,013	0,016	0,017	...	0,01831

Для визначення числових характеристик розглянемо достатньо великий інтервал часу T . Протягом цього часу через систему пройде в серед-

ньому λT вимог. Середнє значення тривалості сумарного часу обслуговування λT вимог дорівнює:

$$\frac{S}{\mu_0} \lambda T = S \psi T .$$

Отже, якщо замість S каналів, що функціонують протягом інтервалу T , розглядати лише один канал з параметром μ_0/S , який обслуговує вимоги в цьому ж інтервалі часу T , то повний час зайнятості цього каналу буде ψST і відповідно повний час незайнятості – $(1-\psi)ST$. При цьому можна вважати, що в системі, що розглядається, є в середньому $\bar{j} = \psi S$ зайнятих каналів, а середня кількість вільних каналів

$$\bar{r} = (1-\psi)S . \quad (2.31)$$

З іншого боку, середній час очікування усіх вимог, що знаходяться в черзі, дорівнює $\lambda T \bar{\omega}$, звідки середня кількість вимог у черзі

$$\bar{v} = \frac{\lambda T \bar{\omega}}{T} = \lambda \bar{\omega} . \quad (2.32)$$

Аналогічно для середньої кількості вимог у системі маємо:

$$\bar{n} = \lambda \bar{u} . \quad (2.33)$$

Вирази (2.31) і (2.32) називаються формулами Літгла. На основі наведених міркувань можна знайти числові характеристики системи $M/M/S/\infty$. Для середньої кількості вимог у черзі можна записати:

$$v = \sum_{n=S+1}^{\infty} (n-S) P_n = \sum_{n=S+1}^{\infty} (n-S) \frac{S^S \psi^n}{S!} P_0 . \quad (2.34)$$

Остаточний розв'язок (2.34) має вигляд:

$$\bar{v} = \frac{S^S \psi^{S+1}}{S!(1-\psi)^2} P_0 . \quad (2.35)$$

Використовуючи співвідношення (2.31)-(2.35), знаходимо вираз для середньої кількості вимог у системі:

$$\bar{n} = \bar{v} + \bar{j} = \bar{v} + S - \bar{r} = \frac{S^S \psi^{S+1}}{S!(1-\psi)^2} P_0 + S \psi .$$

За допомогою формул Літтла одержуємо:

$$\bar{\omega} = \frac{P_0}{\mu} \frac{S^S \psi^S}{S!(1-\psi)^2},$$

$$\bar{u} = \frac{\bar{n}}{\lambda} = \frac{1}{\mu} \left[\frac{S^S \psi^S}{S!(1-\psi)^2} P_0 + S \right].$$

Імовірність того, що в черзі очікує принаймні одна вимога, дорівнює імовірності того, що в момент надходження вимоги в системі знаходиться більш ніж $(S - 1)$ вимог:

$$P(\omega > 0) = P(N \geq S) = \sum_{n=S}^{\infty} P_n = \frac{S^S \psi^S}{S!(1-\psi)} P_0.$$

2.2.2 Багатоканальна СМО з неоднаковими приладами

У випадку неоднакових приладів повна інтенсивність обслуговування визначається сумою $\mu_0 = \sum_{i=1}^S \mu_i$, яка характеризує середню кількість вимог, що можуть обслуговуватися при постійній зайнятості приладів за одиницю часу.

Оскільки розв'язання задачі з S приладами пов'язане з дуже громіздкими обчисленнями, наведемо лише основні характеристики для випадку $S=2$. При цьому припустимо, що вимога, яка надійшла в момент відсутності черги, може надійти на обслуговування в кожному з двох каналів з імовірністю φ_i . Встановимо

$$\varphi_1 = \varphi, \varphi_2 = 1 - \varphi, 0 \leq \varphi \leq 1. \quad (2.36)$$

Припустимо, що канал 1 обслуговує швидше каналу 2, тобто $\mu_1 \geq \mu_2$. У розглянутому випадку якість обслуговування залежить від чотирьох параметрів λ , μ_1 , μ_2 і φ . Але якщо розглядати коефіцієнт використання системи

$$\psi = \frac{\lambda}{\mu_0} = \frac{\lambda}{\mu_1 + \mu_2}$$

і відношення інтенсивностей

$$\alpha = \frac{\mu_2}{\mu_1},$$

то ефективність системи буде залежати тільки від цих двох параметрів ψ і α , причому величина α припускається меншою одиниці.

Для сталого режиму за допомогою диференціального методу одержують рівняння стану:

$$\begin{aligned}\lambda P_0 &= \mu_2 P_{0,1} + \mu_1 P_{1,0}; \\ (\lambda + \mu_2) P_{0,1} &= \mu_1 P_2 + (1 - \phi) \lambda P_0; \\ (\lambda + \mu_1) P_{1,0} &= \mu_2 P_2 + \phi \lambda P_0; \\ (\lambda + \mu_0) P_n &= \mu P_{n+1} + \lambda P_{n-1}, \quad n \geq 2,\end{aligned}$$

де $P_{1,0}$ – імовірність того, що при $n=1$ зайнято канал 1, а другий вільний;

$P_{0,1}$ – імовірність того, що при $n=1$ зайнято канал 2, а перший вільний.

Відповідні імовірності:

$$\begin{aligned}P_{0,1} &= \frac{\psi}{1 + 2\psi} \frac{1 + \alpha}{\alpha} (\psi + 1 - \phi) P_0; \\ P_{1,0} &= \frac{\psi}{1 + 2\psi} (1 + \alpha) (\psi + \phi) P_0; \\ P_n &= \frac{\psi^n}{1 + 2\psi} \frac{1 + \alpha}{\alpha} [1 + (1 + \alpha)\psi - (1 - \alpha)\phi] P_0.\end{aligned}$$

Середня кількість вимог у системі

$$\bar{n} = \frac{1 - P_0}{1 - \psi},$$

де

$$1 - P_0 = \psi(1 + \alpha) \frac{1 + (1 + \alpha)\psi - (1 - \alpha)\phi}{\alpha(1 + 2\psi) + \psi[1 + (1 + \alpha^2)\psi - (1 - \alpha^2)\phi]}.$$

Знаючи $\bar{n}$, можна легко знайти всі інші характеристики якості роботи системи.

2.3 Замкнені СМО

Як уже відзначалося, для замкнених СМО є характерною обмежена і постійна кількість джерел вимог. Прикладом таких систем є обслуговування персоналом деяких пристроїв (верстатів), що можуть відновлюватися після відмов; пристрої є “джерелами вимог”, а персонал –

“обслуговуючими приладами”. Особливістю замкнених СМО є те, що кожне джерело не може посилати наступну вимогу доти, доки послана ним раніше вимога не була обслугована (скажімо, пристрій відновлений і продовжує роботу до нової відмови). Така особливість позначається на інтенсивності потоку вимог, що надходять на прилади системи.

Розглянемо замкнену СМО при таких припущеннях. Маємо m джерел вимог і S обслуговуючих приладів, при цьому $m > S$. Кожне джерело створює найпростіший потік з параметром λ . Інтенсивність обслуговування вимог кожним приладом дорівнює μ . Нехай n – кількість вимог, що надходять від джерел, λ_n – сумарна інтенсивність потоку, тоді початковими параметрами для замкненої СМО є:

$$\left. \begin{array}{l} \lambda_n = m\lambda \\ \mu_n = 0 \end{array} \right\} \text{ при } n = 0, \quad \left. \begin{array}{l} \lambda_n = (m-n)\lambda \\ \mu_n = n\mu \end{array} \right\} \text{ при } 1 \leq n < S,$$

$$\left. \begin{array}{l} \lambda_n = (m-n)\lambda \\ \mu_n = S\mu \end{array} \right\} \text{ при } S \leq n \leq m.$$

Для сталого режиму рівняння станів мають вигляд:

$$\begin{aligned} m\lambda P_0 &= \mu P_1, \\ (m-n)\lambda + n\mu P_n &= (m-n+1)\lambda P_{n-1} + (n+1)\mu P_{n+1}, \quad 1 \leq n \leq S, \\ (m-n)\lambda + S\mu P_n &= (m-n+1)\lambda P_{n-1} + S\mu P_{n+1}, \quad S \leq n \leq m, \\ S\mu P_m &= \lambda P_{m-1}. \end{aligned}$$

За допомогою цих рівнянь одержимо такі рекурентні формули:

$$\begin{aligned} P_n &= \frac{m-n+1}{n} \psi P_{n-1}, \quad 0 \leq n \leq S, \\ P_n &= \frac{m-n+1}{S} \psi P_{n-1}, \quad S \leq n \leq m, \end{aligned}$$

які приводяться до більш зручного для розрахунків вигляду:

$$\begin{aligned} P_n &= C_m^n \psi^n P_0, \quad 0 \leq n \leq S, \\ P_n &= \frac{n!}{S! S^{n-S}} C_m^n \psi^n P_0, \quad S \leq n \leq m. \end{aligned}$$

де

$$C_m^n = \frac{m!}{(m-n)!n!}, \quad \varphi = \frac{\lambda}{\mu S}, \quad p_0 = \frac{1}{\sum_{n=0}^{S-1} C_m^n S^n \psi^n + \sum_{n=S}^m C_m^n \frac{n! S^S}{S!} \psi^n}.$$

Середня кількість вимог у черзі і в системі визначається виразами:

$$\bar{v} = \sum_{n=S+1}^m (n-S)P_n = \frac{S^S P_n}{S!} \sum_{n=S+1}^m \frac{(n-S)n!}{S_n} C_m^n \psi^n,$$

$$\bar{n} = \sum_{n=0}^m nP_n = \sum_{n=0}^S nP_n + \sum_{n=S+1}^m nP_n = \bar{v} + S - \bar{r},$$

де $\bar{r}$ – середнє значення незайнятих каналів: $\bar{r} = \sum_{n=0}^S (S-n)P_n$.

Для визначення середнього часу очікування вимогами початку обслуговування може бути використана залежність:

$$\bar{\omega} = \frac{\bar{v}}{\mu(S - \bar{r})}.$$

Таким чином, отримано всі необхідні формули для розрахунку замкнених СМО.

2.4 Багатофазні СМО

Багатофазні СМО – найбільш розповсюджені системи для формалізації різних виробничих процесів і технічних систем. Їх аналіз є більш трудомістким і складним в порівнянні з розглянутими СМО. Складність аналізу залежить від багатьох факторів, до числа яких відносяться: обмеження довжини v перед фазами, закони розподілу тривалості обслуговування вимог тощо. Розглянемо двофазну СМО, на вхід якої надходить пуассонівський потік з параметром λ , час обслуговування вимог розподілено за експоненціальним законом з параметром μ_1 для першої фази і параметром μ_2 – для другої. У кожній фазі є один обслуговуючий канал, перед яким може утворюватися черга необмеженої довжини.

Припустимо, що вимоги в системі не втрачаються. Тоді і для другої фази параметр вхідного потоку буде дорівнює λ . Втрата вимог може бути врахована відповідним зменшенням λ .

Позначимо через $P(n_1, n_2, t)$ імовірність того, що в момент t у першій фазі є n_1 вимог, а в другій фазі – n_2 вимог. Рівняння стану двофазної системи мають вигляд:

$$\frac{dP(0,0,t)}{dt} = -\lambda P(0,0,t) + \mu_2 P(0,0,t), \quad n_1 = n_2 = 0;$$

$$\frac{dP(0,n_2,t)}{dt} = -(\lambda + \mu_2)P(0,n_2,t) + \mu_1 P(1,n_2 - 1,t) + \mu_2 P(0,n_2 + 1,t),$$

$$n_1 = 0, \quad n_2 > 0;$$

$$\frac{dP(n_1,0,t)}{dt} = -(\lambda + \mu_1)P(n_1,0,t) + \mu_2 P(n_1,1,t) + \lambda P(n_1 - 1,0,t),$$

$$n_1 > 0, \quad n_2 = 0;$$

$$\frac{dP(n_1,n_2,t)}{dt} = -(\lambda + \mu_1 + \mu_2)P(n_1,n_2,t) + \mu_1 P(n_1 + 1,n_2,t) +$$

$$+ \mu_2 P(n_1,n_2 + 1,t) + \lambda P(n_1 - 1,n_2,t), \quad n_1 > 0, \quad n_2 > 0.$$

Розв'язком цих рівнянь для стаціонарного стану є вираз:

$$P(n_1,n_2) = \psi_1 \psi_2 P(0,0),$$

$$\text{де } \psi_1 = \frac{\lambda}{\mu_1}, \quad \psi_2 = \frac{\lambda}{\mu_2}.$$

Оскільки

$$\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} P(n_1,n_2) = 1,$$

то

$$P(0,0) = (1 - \psi_1)(1 - \psi_2).$$

Математичне сподівання кількості вимог, що знаходяться в двофазній системі, можна знайти множенням $P(n_1,n_2)$ на $(n_1 + n_2)$ і підсумовуванням спочатку за n_1 , а потім за n_2 . При цьому одержимо:

$$\bar{n} = \bar{n}_1 + \bar{n}_2 = \frac{\psi_1}{1 - \psi_1} + \frac{\psi_2}{1 - \psi_2},$$

де $\bar{n}_1, \bar{n}_2$ – середня кількість вимог, що знаходяться відповідно в першій і другій фазі.

Середня кількість вимог, що знаходяться в черзі, визначається за формулою:

$$\bar{v} = \bar{v}_1 + \bar{v}_2 = \frac{\psi_1^2}{1 - \psi_1} + \frac{\psi_2^2}{1 - \psi_2}.$$

На основі зроблених припущень про закони розподілу потоків вимог і часу їх обслуговування отримані результати можна узагальнити на довільну кількість фаз r в системі. Тоді розглянуті характеристики можуть бути знайдені за допомогою виразів:

$$\bar{n} = \sum_{i=1}^r \frac{\psi_i}{1 - \psi_i}, \quad \psi_i = \frac{\lambda}{\mu_i}, \quad i = \overline{1, r}; \quad (2.37)$$

$$\bar{v} = \sum_{i=1}^r \frac{\psi_i^2}{1 - \psi_i}. \quad (2.38)$$

Застосовуючи до кожної фази формули Літтла, можна легко визначити за допомогою (2.37) і (2.38) часові характеристики багатофазної СМО $\bar{u}$ та $\bar{w}$.

2.5 Основні дисципліни обслуговування вимог вхідних потоків

2.5.1 Загальні положення

Вимоги, що мають перевагу (пріоритет) відносно інших, розрізняються за типом переваг. Розрізняють відносні й абсолютні пріоритети, які характеризують дисципліну обслуговування вимог, що надходять.

Вимоги, що мають відносний пріоритет, при надходженні в систему не переривають вже початого обслуговування вимог, у яких немає переваг або є такі самі. Вимога з відносним пріоритетом починає обслуговуватися відразу після закінчення обслуговування вимоги без переваги, якщо вона знаходилася в приладі в момент надходження вимоги з пріоритетом.

Якщо вимозі надано абсолютного пріоритету, то вона при надходженні в систему перериває вже почате обслуговування вимог з відносним пріоритетом чи без нього. Вимога, обслуговування якої було перервано, повертається в чергу і очікує дообслуговування чи повного обслуговування.

Встановлена дисципліна обслуговування вимог впливає на значення характеристик якості роботи СМО і її економічні показники.

2.5.2 Вплив дисципліни з відносним пріоритетом на час очікування в черзі

Для визначення впливу дисципліни обслуговування введемо такі позначення:

r_i – ознака переваги: чим менше i ($i = 1, 2, 3, \dots, K$), тим більшу перевагу має вимога і тим раніше вона надходить з черги в обслуговуючий канал; вимоги, що мають ознаку r_K , не мають переваг;

μ_{ri} – інтенсивність обслуговування вимоги з ознакою r_i ;

λ_{ri} – інтенсивність потоків вимог з ознакою r_i ;

$\lambda = \sum_{i=1}^K \lambda_{ri}$ – сумарний потік вимог;

$\psi_{ri} = \lambda_{ri} / \mu_{ri}$ – завантаження приладу системи потоком вимог з ознакою r_i ;

$\psi_r = \sum_{i=1}^K \psi_{ri}$ – сумарне завантаження приладу всіма потоками вимог;

$\psi_{r-1} = \sum_{i=1}^{K-1} \psi_{ri}$ – завантаження приладу вимогами, що мають більший пріоритет у порівнянні з вимогами r_K -пріоритету.

Середній час очікування в черзі для вимог з пріоритетом r_K має вигляд:

$$\bar{\omega}_{r_K} = \frac{\lambda}{2} \frac{\sigma_t^2 + \bar{t}_{об}^2}{(1 - \psi_{r-1})(1 - \psi_r)}, \quad (2.39)$$

де σ_t^2 – дисперсія часу обслуговування;

$\bar{t}_{об} = \frac{1}{\mu}$ – середній час обслуговування довільної вимоги.

Середній час очікування в черзі для вимог, що мають пріоритет $r - 1$, буде меншим, ніж розрахований за формулою (2.39). Він визначається виразом

$$\bar{\omega}_{r-1} = \frac{\lambda}{2} \frac{\sigma_t^2 + \bar{t}_{об}^2}{1 - \psi_{r-1}}. \quad (2.40)$$

Розглянемо окремий випадок, коли середній час обслуговування однаковий для вимог усіх пріоритетів. Можна показати, що при цьому отримуємо вираз:

$$\bar{\omega} = \frac{\lambda}{2} \frac{\sigma_t^2 + \bar{t}_{об}^2}{1 - \psi},$$

з якого видно, що середній час очікування в черзі не залежить від наданих вимогам пріоритетів і дорівнює середньому часу очікування для випадку справедливої дисципліни обслуговування (першим прийшов – перший обслугований). Ефект переваги при цьому полягає в зменшенні часу очікування вимог з великою перевагою за рахунок вимог з меншою перевагою.

Тепер припустимо, що є дві категорії вимог з індексами r_1 і r_2 , для яких одержуємо:

$$\bar{\omega}_{r_1} = \frac{\alpha}{1 - \psi_{r_1}}, \quad \bar{\omega}_{r_2} = \frac{\alpha}{(1 - \psi_{r_2})(1 - \psi)},$$

де

$$\alpha = \frac{\lambda}{2} (\sigma_t^2 + \bar{t}_{об}^2), \quad \psi_{r_1} = \frac{\lambda_1}{\mu_{r_1}}, \quad \psi_{r_2} = \frac{\lambda_1}{\mu_{r_2}}, \quad \psi = \psi_{r_1} + \psi_{r_2}.$$

Середня тривалість очікування для сукупності вимог

$$\bar{\omega} = \frac{\lambda_1}{\lambda} \bar{\omega}_{r_1} + \frac{\lambda_2}{\lambda} \bar{\omega}_{r_2} = \frac{1 - \beta\psi}{1 - \psi_{r_1}} \frac{\alpha}{1 - \psi},$$

де $\beta = \frac{\lambda_1}{\lambda}$.

Якщо $(1 - \beta\psi)/(1 - \psi_{r_1}) < 1$, то $\bar{\omega} < \alpha/(1 - \psi)$. Можна показати, що це досягається при $\mu_{r_2} < \mu_{r_1}$. Отже, якщо перевагу r_1 присвоїти вимогам з меншою тривалістю обслуговування, то цим досягнемо кращого функціонування системи (одержуємо менше значення $\bar{\omega}$).

Однак різні вимоги можуть відрізнятися різною величиною збитку від їх простою в черзі, що може характеризуватися вартістю простою C_{r_i} .

В цьому випадку необхідно розглядати економічну функцію, яка для двох категорій вимог буде мати вигляд:

$$\gamma = C_{r_1} \lambda_1 \bar{\omega}_{r_1} + C_{r_2} \lambda_2 \bar{\omega}_{r_2}.$$

Нескладно довести, що переваги потрібно надавати тим вимогам, для яких добуток $C_{r_i} \mu_{r_i}$ буде більшим. При введенні такої дисципліни обслуговування виходить менше значення економічної функції γ .

Питання для самоконтролю

1. Сформулюйте сутність диференціального методу аналізу СМО.
2. Запишіть математичну модель системи $M/M/1/\infty$ для сталого режиму її роботи.
3. Наведіть графік залежності $\bar{\omega} = f(\psi)$ і поясніть його відмінність від відповідного графіка $\bar{u} = f(\psi)$.
4. Сформулюйте сутність інтегрального методу аналізу СМО; вкажіть, для яких цілей він застосовується.
5. Навіщо потрібно знати закон розподілу вимог, що виходять із системи? Наведіть його математичну модель для загального випадку.
6. У якому випадку двоканальна СМО, кожен канал якої характеризується інтенсивністю обслуговування μ , краща за одноканальну СМО, що має інтенсивність обслуговування 2μ ? Проаналізуйте при цьому всі характеристики якості роботи таких систем, задавши довільне значення для λ і μ .
7. У чому відмінність аналізу багатоканальних СМО з однаковими і неоднаковими каналами?
8. Порівняйте вирази для характеристик якості роботи замкнених і розімкнених СМО. Для якої з цих СМО трудомісткість обчислення характеристик більша і чому?
9. Запишіть вирази для середнього часу перебування вимоги в багатозафазній системі та у чергах до кожної фази.
10. Наведіть приклади різних дисциплін обслуговування вимог з життя і виробництва.
11. Є два види вимог, для однієї з яких встановлюється спочатку відносний, а потім абсолютний пріоритет обслуговування. Поясніть, у якому з цих випадків час очікування для вимоги без пріоритету буде меншим і чому?

3 НЕМАРКОВСЬКІ МОДЕЛІ МАСОВОГО ОБСЛУГОВУВАННЯ

До немарковських (непуасонівських) систем масового обслуговування відносяться ті, для яких або час обслуговування вимог розподілено не за експоненціальним законом, або вхідний потік є непуасонівським, або і те і інше разом. Розглянемо системи, у яких час обслуговування має не показниковий розподіл, а вхідний потік є пуасонівським.

3.1 Система $M/G/1/\infty$

У загальному випадку такі системи вже були розглянуті в попередньому розділі. В підпункті 2.1.3. були наведені формули для функцій розподілу часу очікування в черзі і системі, які записані в перетвореннях Карсона-Лапласа:

$$F(p) = (1 - \psi) \frac{p}{p - \lambda S'(p)}, \quad (3.1)$$

$$G(p) = [1 - S'(p)]F(p), \quad (3.2)$$

у яких через $S'(p)$ позначено довільний закон часу обслуговування вимог.

У зв'язку з цим за допомогою (3.1) і (3.2) для відомого закону $S'(p)$ можна одержати усі числові характеристики якості роботи непуасонівської СМО розглянутого класу. Покажемо це одразу для двох законів розподілу, які широко застосовуються в інженерній практиці.

3.1.1 Система $M/E/1/\infty$

Для ерлангенського розподілу функція щільності часу обслуговування записується у вигляді:

$$S(t) = \frac{(\mu k)^k}{(k-1)!} e^{-\mu k t} t^{k-1},$$

перетворення Карсона-Лапласа якої має вигляд:

$$S(p) = p \left(\frac{\mu k}{p + \mu k} \right)^k,$$

а

$$S'(p) = 1 - \left(\frac{\mu k}{p + \mu k} \right)^k. \quad (3.3)$$

Підставивши (3.3) у (3.1), знайдемо:

$$F(p) = \frac{(1-\psi)p}{p - \lambda + \lambda \left(\frac{\mu k}{p + \mu k} \right)} k. \quad (3.4)$$

Використовуючи (3.4), можна знайти його оригінал за допомогою таблиць перетворень Лапласа, а потім визначити середній час перебування вимог у черзі $\bar{\omega}$.

Однак знайти $\bar{\omega}$ можна значно простіше, якщо скористатися формулою Поллачека-Хінчина, отриманою для системи $M/G/1/\infty$. Ця формула (наведена раніше) має вигляд:

$$\omega = \frac{\lambda}{2} \frac{\sigma_t^2 + t_{об}^{-2}}{1 - \psi}. \quad (3.5)$$

Для ерлангенського розподілу маємо:

$$t_{об}^{-2} = \frac{k}{\mu}, \quad (3.6)$$

$$\sigma_t^2 = \frac{t}{k\mu^2}. \quad (3.7)$$

Підставивши (3.6) і (3.7) у (3.5), одержимо:

$$\mu \bar{\omega} = \frac{\psi}{2(1-\psi)} \frac{k+1}{k}. \quad (3.8)$$

Шляхом розкладання (3.8) у ряд і виконання нескладних перетворень можна знайти дисперсію часу очікування в черзі, яка визначається виразом:

$$\mu^2 \sigma_{\omega}^2 = \frac{\psi}{1-\psi} \frac{k+1}{k} \left(\frac{\psi}{1-\psi} \frac{k+1}{4k} + \frac{k+2}{3k} \right). \quad (3.9)$$

Вирази (3.8) і (3.9) можуть бути використані для оцінювання часу очікування в черзі і його дисперсії при інших законах розподілу, якщо згадати умови переходу ерлангенського розподілу в інші закони розподілу.

3.1.2 Система $M/D/1/\infty$

Для цієї системи характерна сталість (регулярність) часу обслуговування вимог потоку, що надходить. Припустимо $\overline{t_{об}} = \frac{1}{\mu} = const$ і скористаємося формулою Поллачека-Хінчина.

Оскільки час обслуговування усіх вимог не є випадковою величиною, його дисперсія $\sigma_t^2 = 0$. При цьому з (3.5) маємо:

$$\overline{\omega} = \frac{\lambda}{2} \frac{t_{об}^2}{1 - \psi} = \frac{\psi}{2\mu(1 - \psi)}. \quad (3.10)$$

Порівнявши (3.10) з аналогічним виразом для системи $M/M/1/\infty$, можна зробити висновок: при однаковому значенні часу обслуговування (у середньому) для системи $M/D/1/\infty$ час очікування в черзі виходить у 2 рази меншим, ніж при експоненціальному розподілі часу обслуговування. Зауважимо, що результат (3.10) можна одержати з (3.8), якщо припустити $k = \infty$.

3.2 Обслуговування неоднорідних потоків вимог

Нагадаємо, що під неоднорідним потоком розуміється кілька потоків різних або однакових за фізичною природою і призначенням вимог, які вимагають для себе в середньому різного часу обслуговування. Прикладами таких потоків можуть служити потоки різної за конструкцією і призначенням техніки, що надходить на ремонтне підприємство; обслуговування оператором різних за складністю технічних пристроїв; повідомлення різної довжини, що циркулюють в АСУ між засобами та об'єктами керування тощо.

Припустимо, що в одноканальну систему масового обслуговування надходять m найпростіших неоднорідних потоків вимог з параметрами λ_i ($i = \overline{1, m}$). Час обслуговування вимог кожного потоку розподілено за експоненціальним законом з параметром μ_i . Таке подання неоднорідних потоків дозволяє спростити дослідження, в залежності від конкретного змісту операцій обслуговування, конкретної фізичної природи і призначення вимог.

Припустимо, що існує сталий режим роботи, при якому система статистично стійка.

Позначимо інтенсивність завантаження приладу від кожного потоку $\psi_i = \lambda_i / \mu_i$, тоді сумарна інтенсивність $\psi_0 = \sum_{i=1}^m \psi_i$ повинна бути менше

одиниці. Тільки за цієї умови система може працювати в стаціонарному (стійкому) режимі.

Для аналізу розглянутої системи необхідно визначити закон розподілу і параметр вхідного потоку. Оскільки кожен потік (згідно зі зробленим припущенням) є найпростішим, то і сумарний потік буде найпростішим (пуасонівським) з параметром $\lambda_0 = \sum_{i=1}^m \lambda_i$.

Для одержання характеристик якості роботи скористаємося інтегральним методом аналізу, розглянутим у підрозділі 2.1.3. При цьому обмежимося лише двома неоднорідними потоками, тобто припустимо $m=2$, для яких вираз щільності розподілу часу обслуговування запишемо у вигляді:

$$S(t) = \frac{\lambda_1}{\lambda_0} \mu_1 e^{-\mu_1 t} + \frac{\lambda_2}{\lambda_0} \mu_2 e^{-\mu_2 t},$$

а перетворення Карсона-Лапласа $S'(t)$ буде

$$S'(t) = \frac{\lambda_1}{\lambda_0} \frac{p}{p + \mu_1} + \frac{\lambda_2}{\lambda_0} \frac{p}{p + \mu_2}. \quad (3.11)$$

Підставивши (3.11) у (2.22) і виконавши необхідні перетворення, знаходимо:

$$F(p) = F(0) \left[1 + \frac{p\lambda_0}{(p-a)(p-b)} + \frac{\lambda_1\mu_2 + \lambda_2\mu_1}{(p-a)(p-b)} \right], \quad (3.12)$$

де

$$a = -\frac{(\mu_1 + \mu_2 - \lambda_0) - \sqrt{\lambda_0^2 - 2(\lambda_1 - \lambda_2)(\mu_1 - \mu_2) + (\mu_1 - \mu_2)^2}}{2},$$

$$b = -\frac{(\mu_1 + \mu_2 - \lambda_0) + \sqrt{\lambda_0^2 - 2(\lambda_1 - \lambda_2)(\mu_1 - \mu_2) + (\mu_1 - \mu_2)^2}}{2} - \text{корені}$$

знаменника $p - \lambda_0$ $S''(p)$ рівняння (2.22). Виразу (3.12) відповідає оригінал:

$$F(p) = F(0) \left[1 + \lambda_0 \frac{e^{at} - e^{bt}}{a - b} + (\lambda_1\mu_2 + \lambda_2\mu_1) \left(\frac{1}{ab} + \frac{be^{at} - ae^{bt}}{ab(a-b)} \right) \right]. \quad (3.13)$$

Для виконання умови існування сталого режиму необхідно, щоб $\psi_0 = \psi_1 + \psi_2 < 1$. Тоді при $t \rightarrow \infty$ повинна виконуватися умова $F(t) \rightarrow 1$. Оскільки $P(\omega < \infty) = 1$, одержуємо:

$$\lim_{x \rightarrow \infty} F(t) = F(0) \frac{1}{1 - \psi_0} = 1. \quad (3.14)$$

З (3.13) знаходимо:

$$F(0) = 1 - \psi_1 - \psi_2 = 1 - \psi_0, \quad (3.15)$$

що характеризує імовірність відсутності вимог у системі: $F(0) = P_0$.

Підставивши (3.15) у (3.13), остаточно одержимо:

$$F(t) = t - c_1 e^{bt} + c_2 e^{At}, \quad (3.16)$$

де

$$c_1 = \frac{\lambda_0}{A} - \frac{\psi_0(\mu_1 + \mu_2 + \lambda_0)}{2A} + \frac{\psi_0}{2},$$

$$c_2 = c_1 - \psi_0, \quad A = \sqrt{\lambda_0^2 - 2(\lambda_1 - \lambda_2)(\mu_1 - \mu_2) + (\mu_1 - \mu_2)^2}.$$

За допомогою (3.16) можна одержати вираз для середнього часу очікування в черзі довільної вимоги:

$$\bar{\omega} = \int_0^{\infty} t dF(t). \quad (3.17)$$

Провівши інтегрування (3.17), знаходимо: $\bar{\omega} = \frac{\frac{\psi_1}{\mu_1} + \frac{\psi_2}{\mu_2}}{1 - \psi_0}.$

Виконавши аналогічні перетворення виразу (2.20), знайдемо оригінал функції розподілу часу перебування вимог у системі, що має вигляд:

$$G(t) = 1 - e^{bt} \left[c_1 - \frac{\lambda_1 c_1 b}{\lambda_0 (b - \mu_1)} - \frac{\lambda_2 c_1 b}{\lambda_0 (b - \mu_2)} \right] +$$

$$+ e^{at} \left[c_2 - \frac{\lambda_1 c_2 a}{\lambda_0 (a - \mu_1)} - \frac{\lambda_2 c_2 a}{\lambda_0 (a - \mu_2)} \right] -$$

$$-e^{-\mu_1 t} \frac{\lambda_1}{\lambda_0} \left[1 + \frac{c_1 \mu_1}{b - \mu_1} - \frac{c_2 \mu_1}{a - \mu_1} \right] - e^{-\mu_2 t} \frac{\lambda_2}{\lambda_0} \left[1 + \frac{c_1 \mu_2}{b - \mu_2} - \frac{c_2 \mu_2}{a - \mu_2} \right]. \quad (3.18)$$

Підставивши (3.18) в інтеграл $\bar{u} = \int_0^\infty t dG(t)$, визначимо розрахункову формулу для середнього часу перебування вимог у системі:

$$\bar{u} = \frac{\frac{\psi_1}{\mu_1} + \frac{\psi_2}{\mu_2}}{1 - \psi_0} + \frac{\psi_0}{\lambda_0}.$$

Виконавши аналогічні операції для трьох неоднорідних потоків, знайдемо рекурентні формули, що дозволяють розрахувати $\bar{\omega}$ і $\bar{u}$ для довільного числа потоків:

$$\bar{\omega} = \frac{\sum_{i=1}^m \frac{\psi_i}{\mu_i}}{1 - \psi_0}, \quad (3.19)$$

$$\bar{u} = \bar{\omega} + \frac{\sum_{i=1}^m \psi_i}{\sum_{i=1}^m \lambda_i}. \quad (3.20)$$

Останній член (3.20) характеризує середній час обслуговування довільної вимоги:

$$\bar{t}_{об} = \frac{\sum_{i=1}^m \psi_i}{\sum_{i=1}^m \lambda_i}. \quad (3.21)$$

Маючи (3.19) і (3.20), можна знайти середню кількість вимог у системі і черзі: $\bar{n} = \lambda_0 \bar{u}$, $\bar{v} = \lambda_0 \bar{\omega}$, а також середню кількість вимог i -го потоку:

$$\bar{n}_i = \frac{\lambda_i}{\lambda_0} \bar{n}, \quad \bar{v}_i = \lambda_i \bar{\omega}.$$

Наведені формули дозволяють оцінити оптимальну кількість приладів і їх оптимальні параметри, що забезпечують час перебування, який задається замовником, у системі, що обслуговує неоднорідні потоки вимог.

3.3 Необхідність врахування неоднорідності потоків вимог, що надходять до системи

При аналізі СМО досить часто не враховується неоднорідність вимог, що надходять до системи. При цьому використовують усереднення часу обслуговування за всіма неоднорідними потоками згідно з (3.21). При такому підході з'являються помилки у визначенні значень характеристик якості роботи системи, що може привести до того, що система буде нероботоздатною. Величина помилки залежить від співвідношення середніх значень часу обслуговування вимог неоднорідних потоків. Для визначення величини помилки, що виникає при використанні формул для однорідних потоків в той час, коли потоки насправді неоднорідні, можна розглянути відношення (3.19) і (2.15), що для $m = 2$ приводиться до вигляду:

$$\frac{\bar{\omega}}{\omega_{cp}} = \frac{\psi_1^2 + \psi_2^2 + \lambda_1 / \mu_2 \psi_2 + \lambda_2 / \mu_1 \psi_1}{\psi_1^2 + \psi_2^2 + 2\psi_1 \psi_2}. \quad (3.22)$$

З рівняння (3.22) випливає, що середнім значенням часу обслуговування можна користуватися тільки при $\mu_1 = \mu_2$, тобто коли потоки однорідні.

Можна показати, що помилка, яка має місце при ігноруванні неоднорідності потоків, при розрахунку характеристик системи оцінюється виразом:

$$\Delta = \frac{B + \frac{\mu_1}{\mu_2} + \frac{\mu_2}{\mu_1}}{B + 2},$$

де

$$B = \frac{\psi_1^2 + \psi_2^2}{\psi_1 \psi_2}.$$

Графік залежності Δ від μ_1 / μ_2 показаний на рис. 3.1.

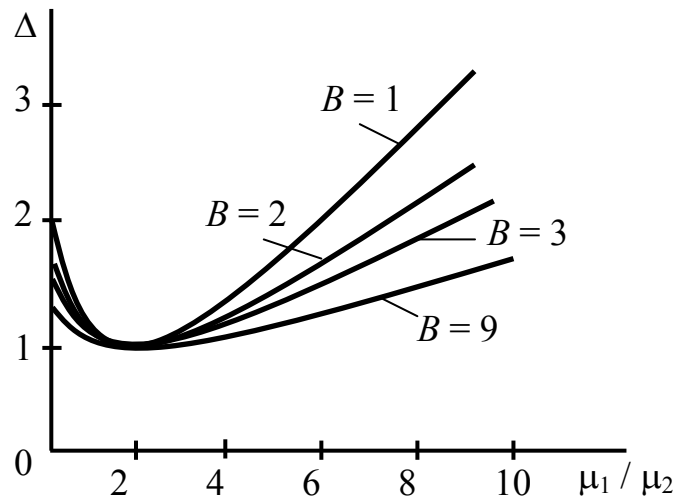


Рисунок 3.1 - Графік залежності Δ від μ_1 / μ_2

З графіка видно, що при $\mu_1 / \mu_2 = 10$ помилка в розрахунках перевищує 100% і не перевищує 3-5%, якщо це відношення знаходиться в межах $0,7 \leq \mu_1 / \mu_2 \leq 1,4$. Останнє можна використовувати для оцінювання неоднорідності потоків: потоки можна вважати однорідними, якщо їх середні значення часу обслуговування вимог відрізняються не більше ніж на 30 %.

Питання для самоконтролю

1. Наведіть за допомогою позначень Кендалла немарковські СМО з відомими законами розподілу потоків вимог і часу їх обслуговування.
2. Запишіть формулу Поллачека-Хінчина і виведіть з неї вираз для середнього часу очікування в черзі вимог з експоненціальним законом часу їх обслуговування.
3. Чому необхідно враховувати неоднорідність потоків вимог?
4. На одноканальну СМО надходить три неоднорідних найпростіших потоки з параметрами $\lambda_1, \lambda_2, \lambda_3$. Тривалості обслуговування вимог цих потоків відрізняються одна від одної більше ніж на 100%. Чи буде на виході системи пуассонівський потік? Поясніть свою відповідь.
5. Знайдіть помилку в розрахунку середнього часу очікування в черзі, якщо використовувати формулу, справедливу для системи $M/M/1/\infty$, в той час як вхідний потік утворений з двох неоднорідних потоків, тривалості обслуговування вимог яких відрізняються в 15 разів.

ЛІТЕРАТУРА

1. Кофман А., Крюон Р. Массовое обслуживание. Теория и приложение: Пер. с фр./ Под ред. И.Н. Коваленко. – М.: Мир, 1965. – 302 с.
2. Саати Т. Элементы теории массового обслуживания и ее приложения: Пер. с англ. / Под ред. И.Н. Коваленко. – М.: Сов. Радио, 1971. – 520 с.
3. Ивченко Г. И., Каштанов В. А., Коваленко И. Н. Теория массового обслуживания. – М.: Высшая школа, 1982. – 156 с.
4. Вентцель Е. С. Исследование операций. – М.: Сов. радио, 1972. – 551 с.
5. Зайченко Ю. П. Исследование операций. – К.: Вища школа, 1975. – 320 с.
6. Таха Хэмди А. Введение в исследование операций: Пер. с англ. – М.: Издательский дом «Вильямс», 2001. – 912 с.
7. Шарейко Л. А. Методические указания к самостоятельной работе. – Винница: ВПИ, 1988. – 60 с.

Навчальне видання

Світлана Борисівна Дубіненко
Юрій Ігорович Мітюшкін

Дослідження операцій
Ч. 1. Системи масового обслуговування

Навчальний посібник

Оригінал-макет підготовлено Мітюшкіним Ю.І.

Редактор О.Д. Скалоцька

Науково-методичний відділ ВНТУ
Свідоцтво Держкомінформу України
серія ДК № 746 від 25.12.2001
21021, м. Вінниця, Хмельницьке шосе, 95, ВНТУ

Підписано до друку
Формат 29,7×42 $\frac{1}{4}$
Друк різнографічний
Тираж прим.
Зам. №

Гарнітура Times New Roman
Папір офсетний
Ум. друк. арк.

Віддруковано в комп'ютерному інформаційно-видавничому центрі
Вінницького національного технічного університету
Свідоцтво Держкомінформу України
серія ДК № 746 від 25.12.2001
21021, м. Вінниця, Хмельницьке шосе, 95, ВНТУ