

МОДЕЛЮВАННЯ ПРОЦЕСІВ КЕРУВАННЯ КОМП'ЮТЕРНИМИ МЕРЕЖАМИ ЗА УМОВ КОНФЛІКТУ

О.П. Ігнатенко

Інститут програмних систем НАН України,
03187, Київ, проспект Академіка Глушкова, 40.
Тел.: +380 (44) 526 3309, факс: 380 (44) 526 62 63, e-mail: ignat@isofts.kiev.ua

Розглядаються проблеми моделювання процесів керування комп'ютерними мережами за умов цілеспрямованих конфліктних дій зі сторони сторонніх гравців.

This work deals with problems of computer networks modelling in situation of denial of service attacks.

Вступ

Сучасні мережі об'єднують незлічені масиви даних, які належать практично до всіх сторін людської діяльності. Захищеність і надійність потоків інформації напряму впливають на якість обслуговування, ефективність роботи і загалом на економічний розвиток цілих галузей виробництва. При цьому складність і взаємопов'язаність мереж постійно зростає. На сьогодні все більше організацій залежать у своїй діяльності від роботи інформаційних мереж. З іншого боку, вразливість мереж також зростає з кожним роком. Таким чином, урахування питань захисту від можливих атак, аналіз роботи при різних аномальних навантаженнях при проектуванні мереж стає нагальною необхідністю.

Особливо важливими і складними для аналізу є інформаційні мережі. Навіть мережа, що моделює роботу невеликого провайдера Інтернет може бути надзвичайно розгалуженою. Проблеми керування, що виникають у мережі Інтернет загалом спільні для всіх типів мереж. Це, зокрема, питання планування і маршрутизації пакетів з одного вузла до іншого. Розглядаючи проблему керування інформаційними мережами важливо зазначити наступні її особливості:

- окремі вузли не мають повної інформації щодо завантаженості буферів та ланок мережі. Всі рішення щодо маршрутизації мають визначатися лише на основі доступної локальної інформації;
- проектні рішення, крім обмеження в інформованості, мають також обмеження у протоколах функціонування мереж;
- у недалекому майбутньому Інтернет буде передавати потоки аудіо– та відеоданих нарівні зі звичайним трафіком, при цьому проблема збалансованого розподілу ресурсів між різнорідними користувачами та забезпечення рівномірної роботи системи залишається невирішеною.

Інтернет–трафік представляє собою надзвичайно складний для аналізу об'єкт, який характеризується періодичністю та нестабільною динамікою. Його дослідженням присвячені чисельні роботи, наприклад, [1, 2]. Часткова причина складності аналізу трафіка полягає у взаємодії великої кількості взаємопов'язаних комп'ютерів, керування яких визначається локальною інформованістю. Недостатність інформації про глобальні характеристики системи виливається у прийняття локально оптимальних рішень. За певних умов такі рішення збігаються з оптимальними, але це трапляється досить рідко. З розвитком мереж з'являються додаткові можливості по інформуванню кожного вузла надійною інформацією про роботу системи за допомогою спеціальних алгоритмів. У зв'язку з цим виникає необхідність у проведенні досліджень і розробці алгоритмів керування потоками даних, які б використовували весь обсяг різноманітної інформації про топологію мережі та рівні завантаженості окремих ланок мережі. Ще одним важливим напрямком розвитку сучасних мереж є безпроводні мережі. На сьогодні вони тільки починають змінювати структуру і функціонування комунікаційних мереж, але необхідність їх подальшого розвитку не викликає сумніву. У безпроводних мережах схеми маршрутизації дуже схожі на ті, що використовуються у мережі Інтернет. Зокрема, в обох випадках ресурсами є живлення і пропускна здатність. Також спільними є наявність багатьох користувачів і можливих шляхів з'єднань між користувачами та вузлами мережі. Відмінність безпроводних мереж полягає у великій мінливості структури в залежності від втрат пакетів та з'єднань, конкуренції користувачів за ресурси системи.

Керування мережами останні п'ятнадцять років було предметом жвавих досліджень, але все ще є досить молодою дисципліною. Основна увага дослідників зосереджена на моделях теорії масового обслуговування, які добре зарекомендували себе при вивченні телефонних ліній та інших моделей мереж. Моделі керування поточковими і стохастичними моделями такого типу розроблені досить повно і надають потужний інструментарій для дослідження стійкості і поведінки систем. Однак суттєвим недоліком таких моделей є, по перше, ідеалізація: для сучасних Інтернет–мереж надзвичайно рідко виконуються базові припущення моделей теорії масового обслуговування (стаціонарність, відсутність післядії). По–друге, в цих моделях не врахована можливість автономного керування (оскільки така можливість для вузлів мережі з'явилась відносно недавно).

По-третє, наявність різних груп користувачів призводить до конкуренції за ресурси системи, тому системи керування необхідно шукати рівноважне рішення, яке б дозволяло не тільки розподіляти ресурси між користувачами але і витримувати певний баланс. Ігрова постановка задачі розподілу ресурсів – один із найбільш популярних напрямків сучасного розвитку безпроводних мереж [3].

Для вирішення задач динамічного планування пропонується застосовувати досягнення теорії конфліктних динамічних систем. Такий підхід, який використовуватиме потужний апарат для опису поведінки керованих систем за умов невизначеності і конфлікту, дозволить розробити аналітичну модель поведінки системи та гарантувати певний рівень обслуговування користувачів за будь яких умов.

1. Моделювання мереж

Виникнення теорії масового обслуговування пов'язують з данським вченим, співробітником телефонної станції А.К. Ерлангом. Його роботи, що стосувалися питань раціонального обслуговування користувачів телефонних станцій, розпочали еру досліджень, які на сьогодні є класичними. Першим предметом розгляду був, природно, найпростіший випадок – системи з одним сервером, які є базовою моделлю теорії масового обслуговування. В нотації Кендалла [4] довільну систему можна записати у вигляді

$$A/B/C/K/N/D$$

де літера *A* описує розподіл вхідного процесу (*M* – марковський, *Ek* – розподіл Ерланга, *G* – загальний розподіл);

B відповідає розподілу часу обслуговування на сервері;

C – кількість серверів у системі;

K – загальна ємність системи (кількість пакетів, що може одночасно в ній перебувати);

N – кількість пакетів або користувачів, що намагаються отримати обслуговування (якщо не вказується, то кількість вважається нескінченною);

D – схема обслуговування пакетів, що надійшли у чергу. Системи з одним вхідним потоком досить добре вивчені на сьогодні (наприклад, [5]).

Мережі черг. Переважна більшість систем масового обслуговування, що зустрічаються в реальному житті, складаються більше ніж з одного обслуговуючого центру. Користувачі або пакети можуть переходити з однієї частини системи в іншу, отримувати послуги у певній послідовності, обслуговуватися повторно тощо. Практично будь-яка складна система може слугувати прикладом мережі черг: технологічні лінії збірки деталей, потоки машин у транспортній мережі міста, системи виконання замовлень, комп'ютерні системи типу «клієнт-сервер», мережа Інтернет, телекомунікаційні мережі, що включають в себе звичайні телефонні лінії та сучасні високо-швидкісні канали, призначені для передачі даних, потокового аудіо- та відео сигналів, інших сервісів.

Перші роботи з дослідження цього напрямку [6] розглядали спеціальні випадки, такі як цикли або лінії з експоненціальних черг. У 1957 р. Джексон [7] довів, що у довільній відкритій мережі, яка складається із серверів з експоненціально розподіленим часом обслуговування та пуасоновською частотою прибуття зовнішніх пакетів, граничний розподіл пакетів буде мати спеціальний вигляд. Розподіл пакетів на *i*-у вузлі буде ідентичний розподілу для окремого сервера типу *M/M/1* з розподілом часу обслуговування таким же, що і на *i*-ому вузлі та розподілом прибуття пакетів, який знаходиться з *рівняння трафіку* і залежить від частот інших вузлів мережі. Це дуже значний результат в теорії масового обслуговування, оскільки відомо, що для довільного вузла розподіл прибуття пакетів не буде пуасоновським (це буде вірно лише для мереж без циклів).

В 70-і роки виявилось, що Джексоновські мережі та подібні до них є потужним інструментом при дослідженні комп'ютерних систем. Ідеї створення комунікаційних мереж, які б об'єднували віддалені комп'ютери в одну систему, та використання досягнень теорії масового обслуговування для їх моделювання особливо обстоювалась Леонардом Клейнроком, який присвятив цій проблемі велику частину своєї роботи [8]. Важливим випадком моделей мереж є моделі зі скінченними обсягами буферів на вузлах та блокуванням. В таких мережах певні вузли мають обмежені можливості за зберіганням пакетів у черзі, тому при перевищенні порогового значення нові пакети можуть бути втрачені або заблоковані. Моделі цього типу часто зустрічаються у реальному житті, але надзвичайно важкі для аналітичних досліджень. Певні результати представлені в [9].

Перевантаження у мережах. Робота сервера сильно впливає на систему коли він перевантажений, тобто довжина черги велика. Таким чином, дослідження максимальних величин черг дає важливі характеристики роботи системи на грані перевантаження. Можливість адекватного моделювання режимів перевантаження має критичне значення при проектуванні мереж. Для складних мереж нерідко єдиним способом дослідження поведінки при таких режимах є імітаційне моделювання. Основні дослідження зосереджуються на пошуках пришвидшення роботи алгоритмів і зменшення помилок при моделюванні. Іншим підходом є пошук підходящих апроксимаційних моделей. Як правило, такі моделі стають все менш надійними при досягненні режимів перевантаження. Дослідження аналізу перевантажень у системах масового обслуговування вперше розглядалось у роботі Кінгмана [10].

Стійкість та потокові моделі. Одним із фундаментальних понять теорії масового обслуговування є поняття *стійкості*. Вважається очевидним, що в разі забезпечення системи достатньою кількістю ресурсів для обробки вхідних пакетів система має з часом виходити на певний режим роботи. Точне значення стійкості визначається для кожної конкретної області застосування окремо, але загальне визначення наступне.

«Стійкість» означає, що характеристики кожного вузла (довжина черги, час затримки) є $o(t)$ при $t \rightarrow \infty$. Або потужність вихідного потоку має бути рівна потужності вхідного. Для стохастичних систем стійкість розуміють як існування певного граничного стаціонарного розподілу пакетів по вузлах мережі, до якого з часом наближається система. Однак за такого підходу з поля зору часто випадають необхідні й достатні умови існування стійкого розв'язку (наприклад, умова $\rho < 1$ для системи $M/M/1$). Альтернативним рішенням є доведення стійкості за допомогою методу функцій Ляпунова, грубо кажучи, знаходиться «потенціальна» функція системи, що монотонно зростає з часом. Якщо ця функція має границю, то система стійка [11].

Неочікуваний і важливий результат у цьому напрямку був отриманий у 1990. Кумаром і Зейдманом [12]. У своїй роботі вони навели приклад потокової моделі мережі, в якій інтенсивність потоку на кожному вузлі менша за одиницю, але система при цьому не є стійкою. Пізніше Рибко і Столяр представили [13] стохастичну систему з ідентичною до детерміністичної моделі Кумара динамікою. Для аналізу стійкості своєї моделі вони використовували потокову модель. Ця ідея виявилась надзвичайно вдалою і породила велику кількість застосувань.

Так, у роботі [14] проводиться узагальнення даної ідеї в контексті неоднорідних мереж черг (multi class queue network MCQN). Зокрема показано, що MCQN стійка (в сенсі існування граничного розподілу) тоді і тільки тоді, коли «відповідна потокова моделі» стійка (в сенсі існування моменту часу, для якого всі черги дорівнюють нулю). Під відповідною потоковою моделлю розуміється система з ідентичною топологією та маршрутизацією, в якій дискретні пакети замінені на неперервні потоки з детерміністичними інтенсивностями, які дорівнюють математичному сподіванню відповідних характеристик початкової стохастичної системи.

Моделі керування. Узагальнення аналізу перевантажень на випадок проблеми керування для MCQN систем було вперше запропоновано в 1988 р. Гаррісоном [15]. Схема аналізу для звичайних мереж включає наступні кроки:

1. Опис моделі MCQN.
2. Побудова ідеалізованої стохастичної мережі (з Гаусовськими розподілами випадкових величин на вузлах).
3. Визначення оптимальної стратегії (policy) для спрощеного випадку.
4. Використання оптимального рішення для отримання «хорошого» відповідного йому рішення для початкової MCQN системи.
5. Доведення збіжності роботи MCQN та ідеалізованої мережі при використанні описаних стратегій.

Ключовий новий крок, який запропонував Гаррісон, полягає у заміні Гаусовської моделі еквівалентною (малорозмірною) моделлю, що описує векторний процес завантаження вузлів системи. Таким чином відбувається зменшення розмірності задачі та спрощення її розв'язку. Під завантаженістю вузла розуміється очікуваний обсяг роботи для цього вузла, що спричиняється кожним пакетом у мережі. З точки зору класичної теорії масового обслуговування задача полягає у пошуку такої функції «роботи» для кожного вузла, яка б описувала, в які періоди часу він працює а в які – ні. Призначення такої функції для кожного вузла та її узгодження з процесом завантаження повністю розв'язує задачу керування.

Цей підхід виявився успішним для розв'язання проблем, в яких оптимальне керування є лінійним. Хоча такий клас задач досить вузький, однак він більший, ніж вихідний клас MCQN систем з лінійним рішенням.

2. Моделювання процесів керування

Сучасні мережі об'єднуються у великі структури, які включають різні типи трафіка (цифрові масиви даних, відео, аудіо, голосові потоки та ін.). Внаслідок виникають нові проблеми, пов'язані, наприклад, з керуванням потоками, динамічним розподілом ресурсів, забезпеченням захисту системи. Крім того, проблеми керування, вже розв'язані для традиційних мереж, набувають принципово іншої складності.

Найбільш поширеними архітектурами телекомунікаційних мереж на сьогодні є Інтернет і АТМ (мережі з асинхронною передачею даних). Інтернет пропонує організацію мережі за принципом рівномірного розподілення ресурсів між усіма користувачами. Якщо кількість користувачів зростає якість послуг (затримки, втрата інформації) для кожного користувача погіршуються. АТМ мережі надають, в основному, гарантований рівень сервісу. Тобто, при завантаженості мережі з'єднання може не відбутися, але якщо з'єднання встановлено, то користувач отримує заявлені послуги не нижче певної якості.

Як зазначалося, основним питанням дослідження поведінки таких мереж є стійкість. За яких умов кількість робіт у чергах зменшується до нуля або принаймі не зростає з часом? При відповіді на це питання неможливо обійти стороною важливий аспект організації роботи мережі – яким чином задані правила вибору наступного пакета з черги для його обробки. Іншими словами, яким чином здійснюється маршрутизація та керування обробкою пакетів.

Очевидно, що існує безліч можливих варіантів. Можливо, найбільш природним є схема FIFO (first-in-first-out), за якої перший пакет, що надійшов на вузол мережі, отримує пріоритет на обробку. Ще одним відомим правилом є PS (processor sharing), коли всі ресурси вузла діляться порівну між усіма наявними роботами (тобто сервер надає частки свого робочого часу всім за чергою).

Правила FIFO та PS називають зрівнюючими в тому розумінні, що пакети в різних чергах або *буферах* отримують рівні умови за обробкою на даному вузлі мережі.

Принципово іншою є схема статичного пріоритету SBP (static buffer priority), за якої кожному буферу задається пріоритет і пакети з вищим пріоритетом обслуговуються у першу чергу.

У межах одного буфера пакети обслуговуються у порядку прибуття. Схема називається випереджувальною, якщо прибуття роботи з вищим пріоритетом перериває обробку поточної роботи з нижчим пріоритетом. Після обробки пріоритетної роботи виконання перерваної поновлюється.

Пошук керувань, які б оптимізували час затримки, витрати на обслуговування мережі та інші характеристики інтенсивно досліджувалось протягом останніх десятиліть. Звичайним підходом до вирішення цих проблем є використання теорії марковських процесів прийняття рішень, який полягає у мінімізації відповідного функціонала. Однак цей підхід дозволяє рішення у явному вигляді лише невеликої кількості задач. Існує, звичайно, можливість розв'язати відповідні рівняння чисельно, за допомогою ітерацій, але, як правило, ситуація суттєво ускладнена наявністю необмеженого простору станів.

Така ситуація спричинила інтерес до дослідження спрощених, ідеалізованих задач керування мережами. Такі моделі завжди використовувалися для дослідження питань стійкості мережі. Однак з'ясувалося, що оптимальне керування потокової моделі та стохастичної мережі надзвичайно схожі [16]. Після виявлення цього факту дослідження були зосереджені на пошуку зв'язків між цими керуваннями. Це дуже важливе питання, тому що оптимальне керування потокової моделі знаходиться в явному вигляді відносно легко [17]. Принаймі, існують ефективні алгоритми вирішення цієї проблеми.

На сьогодні в цій області отримано ряд важливих результатів щодо зв'язку поточкових та стохастичних стратегій керування. Зокрема, Мейном показано [18], що ітеративний процес пошуку правил поведінки процесу сходиться асимптотично до оптимального розв'язку потокової моделі за умови стійкості розв'язку. На жаль, пряме перенесення оптимального керування потокової моделі на стохастичний випадок у цілому не є асимптотично оптимальним.

Проте, як було показано [19], асимптотично оптимальне керування завжди може бути побудовано. Конструкція, запропонована Магларасом, полягає у тому, що при керуванні мережею у певні дискретні моменти часу відбувається переоцінка стану системи і розрахунок стратегії на наступний проміжок часу здійснюється на базі лінійної моделі. Інша ідея висловлена в [16] спирається на той факт, що оптимальне керування потокової моделі є кусково постійною функцією. Певна модифікація поточкового оптимального керування на проміжках часу з постійними значеннями, дає змогу побудувати асимптотично оптимальне керування.

Задача ставить питання про застосування результатів теорії керування.

З точки зору загальної теорії керування можна класифікувати системи організації мереж за об'єктами та цілями керування. Виділяють наступні загальні ситуації.

- Централізоване керування або один контролер. Дана ситуація характерна для задач розміщення заявок: в мережу надходить заявка на нове з'єднання і система має вирішити чи виконувати цю заявку чи залишити її у черзі. Керування полягає у виділенні різним вузлам ресурсів (часу обробки, пропускних каналів тощо).

- Коаліційні ігри. В таких системах існують декілька контролерів зі спільною метою, рішення приймаються агентами, що взаємодіють задля досягнення спільних цілей.

- Антагоністичні ігри. Якщо в системі присутні декілька гравців з автономною поведінкою та власними цілями, що можуть суперечити один одній, то говорять про антагоністичну ситуацію.

- Конфліктні моделі. Антагоністична модель розглядає випадок, коли гравці мають суперечливі цілі, але загалом їх дії спрямовані на отримання послуг від системи. Принципово іншою є ситуація, коли частина гравців об'єднана в групу керованих агентів, які цілеспрямовано намагаються погіршити якість послуг системи.

Керування стохастичними мережами. Нехай стан мережі описується вектором черг $Q(t) = (Q_1(t), \dots, Q_n(t))$, іншими словами поведінка системи визначається завантаженістю n буферів. Визначимо множину індексів $I = \{1, \dots, n\}$, тоді $Q_i(t)$ – обсяг черги з пакетів у i -му буфері у момент часу t .

Буфери Q_i , $i \in I$ об'єднані у мережу за допомогою вузлів. Кожен вузол може містити один або більше буферів та надає їм спільний ресурс для обробки пакетів. Позначимо $S = \{1, \dots, m\}$ множину індексів вузлів мережі. Функція відповідності $s(i) \in S$ визначає приналежність буфера i до вузла $s(i)$.

Матриця відповідності C (*constituency matrix*) визначається за правилом

$$c_{si} = \begin{cases} 1 & s(i) = s \\ 0 & s(i) \neq s \end{cases}, \quad s \in S, i \in I.$$

Після обробки пакет залишає вузол мережі. Може або покинути мережу, або потрапити на інший вузол. Маршрутизація пакетів (яка в рамках даної роботи є детерміністичною) задається матрицею маршрутизації R_0 (*routing matrix*):

$$r_{ij} = \begin{cases} 1 & r(j) = i \\ 0 & r(j) \neq i \end{cases}, \quad j, i \in I,$$

де $r(j) \in I$ – функція, що задає вузол, в який переходять пакети, які залишають вузол j . Знаходження функції керування для такої мережі представляє, взагалі кажучи, непростою задачу. Основними особливостями її є стохастичність і дискретність. Загальна динаміка системи визначається наступним рекурентним співвідношенням:

$$Q(t+1) = Q(t) + B(t+1)U(t) + A(t), \quad (1)$$

де $Q(0) = Q_0$. При цьому вважаються виконаними наступні припущення [4].

1. Будемо казати, що U адаптована до трійки (Q, A, B) , якщо $U(t)$ може бути виражена як функція від випадкових змінних $\{Q(0), \dots, Q(t), A(0), \dots, A(t), B(0), \dots, B(t)\}$ для кожного $t \geq 0$. Також наявні геометричні обмеження

$$U(t) \in \Omega,$$

де $\Omega = \{u \in R^n : u_i \geq 0, Cu \leq \bar{1}\}$, $\bar{1}$ – n -мірний вектор з одиниць, а нерівність виконується порядково.

2. Фазовий вектор процесу (1) також геометрично обмежений:

$$Q(t) \in X \subset R^n.$$

3. $R_0^n = 0_{n \times n}$, тобто у системі немає циклів і кожний пакет може бути оброблений не більше як n раз.

4. $B(t)$ – послідовність незалежних і однаково розподілених матриць, $A(t)$ – послідовність незалежних і однаково розподілених векторів.

При цьому

$$B(t) = -[I - R^T]M(t),$$

де $M(t)$ – послідовність незалежних і однаково розподілених діагональних матриць.

Модель (1) представляє стохастичне дискретне рівняння, розв'язком якого є функція керування $U(t)$. Оскільки маршрутизація вважається фіксованою, то знаходження $U(t)$ рівнозначне розв'язанню задачі планування.

Одне з фундаментальних питань, яке виникає при розв'язанні задачі (1), – за яких умов взагалі існує функція $U(t)$, яка б забезпечувала приведення $Q(t)$ в нуль?

Важливу роль у дослідженнях цього питання відіграють потокові моделі.

Потокові моделі. Позначимо $\alpha = E[A(t)]$, $M = E[M(t)]$, де $M_{ii} = E[M_{ii}(t)]$, тобто M – постійна діагональна матриця. Тоді потокова модель описується простим диференціальним рівнянням

$$\dot{q} = \alpha + Bu(t), \quad (2)$$

де $B = -[I - R^T]M$. Будемо казати, що модель (2) може бути стабілізована, якщо для будь-якого $q(0)$ та функції $\alpha(t)$ існує керування $u(t)$ і час $T \geq 0$, такі, що

$$q(0) + T\alpha + \int_0^T Bu(\tau)d\tau = 0.$$

При цьому мінімальний $T(q(0))$ називається часом вичерпання черги. Пошук мінімального часу для системи (2) – це задача оптимального керування, яка має розв'язок для фіксованих функції $\alpha(t)$ та початкового вектора $q(0)$, якщо виконується включення

$$q(0) - \alpha T \in T \cdot BU. \quad (3)$$

Включення (3) дає можливість визначити момент часу $T(q(0))$. Однак структура потоків керування задачі (2) описується матрицями B і C , що навіть для простих мереж ускладнює знаходження функції керування $u(t)$.

У фундаментальній роботі [4] описаний проблемно-орієнтований підхід, що дозволяє отримати керування для кожного вузла у явному вигляді. Опишемо головну ідею цього підходу, яка полягає у введенні поняття *завантаженості* вузлів мережі.

1. Визначимо *матрицю завантаженості* за допомогою наступного рівняння:

$$\Xi = -CB^{-1} = M^{-1}[I - R^T]^{-1},$$

де M – невідроджена діагональна матриця, а $[I - R^T]^{-1} = \sum_{k=0}^{\infty} R^k = \sum_{k=0}^n R^k$. Позначимо ξ^s , $s \in S$ вектори-рядки, що утворюють матрицю Ξ .

2. *Вектором завантаження* назовемо вектор $\rho = \Xi\alpha$. Введемо позначення $\rho_s = \max_{s \in S} \rho_s$, тоді вузьким місцем називаються такі вектори: ρ_s , що $\rho_s = \rho_s$.

Твердження ([18]).

Модель (2) може бути стабілізована тоді і тільки тоді, коли $\rho_s < 1$. При цьому мінімальний час вичерпання черги

$$T(q) = \max_{s \in S} \frac{(\xi^s, q)}{1 - \rho_s}.$$

Одним із можливих керувань, що забезпечують стабілізацію системи, є вектор

$$u^* = -B^{-1} \left(\alpha + \frac{q(0)}{T(q(0))} \right).$$

Проілюструємо ці результати на модельних прикладах, відомих з теорії черг.

Модель окремого сервера. Потокову модель окремого сервера можна представити у вигляді диференціального рівняння

$$\dot{q} = -\mu u(t) + \alpha.$$

Рівняння має розв'язок при виконанні умови $\mu > \alpha$. Оптимальне (у сенсі задачі швидкодії) керування задається формулою

$$u(t) = \begin{cases} 1, & q(t) > 0, \\ 0, & q(t) = 0. \end{cases}$$

Час вичерпання черги

$$T(q(0)) = \frac{q(0)}{\mu - \alpha}.$$

Моделі Клімова. Узагальненням моделі з одним сервером є ситуація, коли пакети відрізняються за пріоритетом або вимагають суттєво різного часу обробки. В цьому випадку можна вважати, що у системі існують m ліній обробки (можливо віртуальних). Всі пакети, що потрапляють у систему, класифікуються на m типів та опрацьовуються у відповідній лінії. Загальна задача полягає у знаходженні функції керування $u(t) = (u_1(t), \dots, u_m(t))$, яка визначає які ресурси спрямувати у момент часу t на обробку пакетів i -ї лінії на основі інформації про поточне завантаження системи.

Загальна кількість ресурсів обмежена, тому для кожного $t \in [0, T]$

$$\sum_{i=1}^m u_i(t) \leq 1.$$

Розглянемо модель системи у диференціальній формі:

$$\dot{q} = Bu(t) + \alpha.$$

Матриця B має діагональний вигляд: $B = -diag(\mu_1, \dots, \mu_n)$, $C = [1, \dots, 1]$, $\alpha(t) = (\alpha_1, \dots, \alpha_n)$. Модель Клімова для чотирьох черг показана на рис. 1.

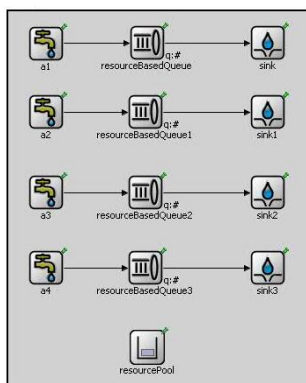


Рис. 1. Модель Клімова для чотирьох черг

Модель Клімова може бути стабілізована, якщо і тільки якщо

$$\sum_{i=1}^n \frac{\alpha_i}{\mu_i} < 1.$$

Мінімальний час вичерпання черги

$$T(q(0)) = \frac{(\xi, q(0))}{1 - \rho} = \frac{\sum_{i=1}^n q_i(0) \mu_i}{1 - \sum_{i=1}^n \frac{\alpha_i}{\mu_i}}.$$

Модель з повторною обробкою. Поточкова модель з повторною обробкою (рис. 2) описується матрицею B наступного вигляду:

$$B = -(I - R^T)M = \begin{pmatrix} -\mu_1 & 0 & 0 \\ \mu_1 & -\mu_2 & 0 \\ 0 & \mu_2 & \mu_3 \end{pmatrix},$$

де $M = \text{diag}(\mu_1, \mu_2, \mu_3)$. При цьому вектор α має вигляд $\alpha = \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \end{pmatrix}$.

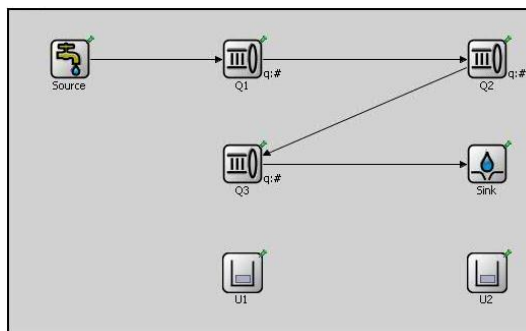


Рис. 2. Модель з повторною обробкою

Використовуючи властивості матриці маршрутизації можна отримати, що

$$B^{-1} = -M^{-1}[I + R^T + R^{2T}].$$

Матриця Ξ складається з двох векторів:

$$\begin{bmatrix} \xi^1 \\ \xi^2 \end{bmatrix} = \begin{bmatrix} \mu_1^{-1} + \mu_3^{-1} & \mu_3^{-1} & \mu_3^{-1} \\ \mu_2^{-1} & \mu_2^{-1} & 0 \end{bmatrix}.$$

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \Xi \alpha = \begin{bmatrix} \frac{\alpha_1}{\mu_1} + \frac{\alpha_1}{\mu_3} \\ \frac{\alpha_1}{\mu_2} \end{bmatrix}.$$

Модель може бути стабілізована тільки у тому разі якщо виконуються нерівності

$$\frac{\alpha_1}{\mu_1} + \frac{\alpha_1}{\mu_3} < 1 \text{ та } \frac{\alpha_1}{\mu_2} < 1.$$

3. Моделювання процесів керування за умов конфлікту

При застосуванні схеми Гаррісона насправді неявно вважаються виконаними багато припущень про роботу мережі. Зокрема вважається, що пакети, які потрапляють на вхід системи, розподілені за законом Пуассона і параметри розподілу не змінюються з часом. Як правило, особливості роботи протоколів (встановлення з'єднання, розриви зв'язку, затримки і переповнення) взагалі не беруться до уваги. Внаслідок дослідження щодо стійкості і завантаженості мережі проводяться для певної «ідеалізованої» моделі. Це, звичайно, дозволяє провести певний аналіз стійкості та отримати інформацію про роботу мережі, але достовірність таких даних має суттєві обмеження. Особливо це стосується ситуацій, коли в роботі мережі виникають аномальні явища. Проблема дослідження аномалій [20], як правило, виникає при побудові систем безпеки. Однак зважаючи на той факт, що до цього поняття належать не тільки вторгнення та атаки на відмову, але і флеш-моби (події, що спричиняють критичне перевантаження системи), то логічно досліджувати динаміку системи за умов таких подій ще на етапі моделювання.

Розглянемо задачу динамічного планування (2). У постановці Мейна [18] ця задача представляє потокову модель з постійним вектором параметрів α і функцією керування $u(t)$. Задача полягає у пошуку функції $u(t)$, такої, що для деякого моменту часу T^* виконується $q(T^*) = 0$ і при цьому функція вартості $c(q(t))$ мінімальна. Розв'язку цієї задачі і пошуку умов на параметри μ_i та α , за яких система буде стійкою, присвячені численні роботи (наприклад, [16 – 19]).

У роботах [21, 22] докладно описувались проблеми конфліктів у мережах та необхідність побудови керування з урахуванням можливих зловмисних дій, спрямованих на перевантаження (атаку на відмову) системи. Одним з важливих кроків у цьому напрямку є введення в мережу додаткового гравця, який намагається вплинути на роботу системи (рис. 3).

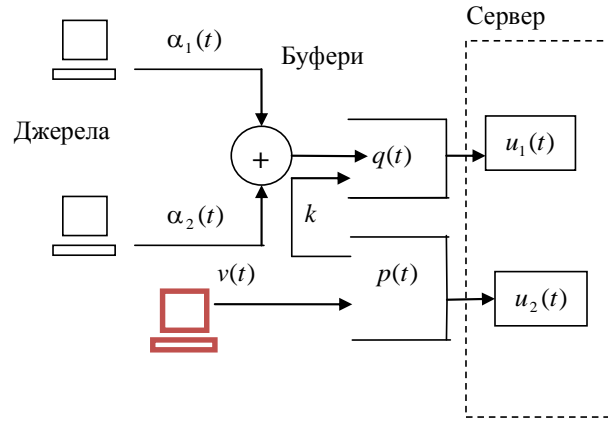


Рис. 3. Система за умов конфлікту

Цей гравець намагається вплинути на роботу системи шляхом вибору функції $v(t)$. Введемо у систему додатковий віртуальний буфер $p(t)$, який буде описувати потужність атаки. Атака відбувається наступним чином: нападник збільшує потужність зі швидкістю $v(t)$, система захисту може відсікати джерела атаки зі швидкістю $u_2(t)$, при цьому, якщо джерело атаки діє, то воно завантажує загальну систему сфальшованими пакетами з коефіцієнтом k .

Будемо вважати величину $p(t)$ – одномірною, $q(t)$, $\alpha(t) = (\alpha_1(t), \dots, \alpha_n(t))$ – n -вимірними векторами.

Розглянемо динамічну систему з початковою умовою $(q(0), p(0))$, яка описується системою лінійних диференціальних рівнянь:

$$\begin{aligned} \dot{q}(t) &= \alpha(t) + k \cdot p(t) \cdot \beta + Bu_1(t), \\ \dot{p}(t) &= v(t) - \mu u_2(t), \quad t \geq 0. \end{aligned} \quad (4)$$

Вектор β має вигляд $\beta = (0, \dots, 1, \dots, 0)$, тобто пакети атаки потрапляють на один з входів мережі. Фазовий стан системи (4) описується вектором $w(t) = (q(t), p(t))$.

Буфери вважаємо обмеженими:

$$0 \leq \|q_i(t)\| \leq q_i^{\max} \quad \text{для всіх } t \geq 0, \quad i \in I.$$

Параметр $p(t) \in \mathcal{R}_+$ пов'язаний з нападником, який намагається погіршити роботу мережі (максимізувати $q(t)$ іншими словами) використовуючи керування $v(t)$, $v(t) \geq 0$, $v(t) \leq v$. Обираючи функцію $v(t)$ в момент часу t нападник встановлює потужність атаки (частоту атакуючих пакетів) $p(t)$. Параметр $p(t)$ впливає на вхідний буфер системи $q(t)$ лінійно з коефіцієнтом $k \geq 0$, вектор β задає буфер, на який потрапляють пакети.

Гра починається з початкової умови $w(0)$. Інший гравець – захисник, розподіляє свої ресурси керування за двома напрямками $u_1(t)$ – для обробки поточних запитів користувачів і пакетів атаки (будемо вважати, що вони завантажують мережу однотипними пакетами) та $u_2(t)$ – для виявлення і відсічення джерел атаки.

На керування захисника накладається природне обмеження $u_1(t) \geq 0$, $u_2(t) \geq 0$, $cu_1 \leq 1$, та $\sum_{i=1}^n u_1^i + u_2 \leq 1$.

При цьому вважається, що захисник у момент часу t володіє інформацією про $\alpha_i(t)$, $v(t)$ і $w(t)$.

Функції $\alpha_i(\cdot): R \rightarrow R^n$, які описують потік пакетів від джерел у момент часу t , вважаються додатними і обмеженими числами α_i^{\max} відповідно. Крім цього загальна кількість пакетів обмежена:

$$\int_{t_0}^{\infty} \alpha_i(t) dt \leq \alpha_i^{\text{int}}, \quad i = 1, \dots, N.$$

Задача полягає у знаходженні умов, які б гарантували наведення вектора $q(t)$ в точку $(0,0)$ зі стану q_0 за будь-яких протидій суперника не пізніше ніж за час $T(q_0) < \infty$.

У роботі [22] доведений наступний результат.

Теорема. Нехай для системи (4) виконуються умови $\mu \geq \nu$:

$$q_1(0) + \frac{k}{2} \frac{(q_2(0))^2}{\mu - \nu} + \sum_{i=1}^N \alpha_i^{\text{int}} \leq q_1^{\max}.$$

Тоді існує стратегія $u(t)$, яка гарантує закінчення гри не пізніше моменту часу $T(q(0))$.

У роботах [21, 22] автором проведено порівняльний аналіз сучасних середовищ моделювання комп'ютерних мереж. Найкращі характеристики для моделювання конфліктних процесів у комп'ютерних мережах, серед яких має середовище OMNeT++. Воно представляє собою інтегроване середовище, призначене для створення, налаштування і запуску імітаційних дискретних моделей. Це середовище, розроблене на основі мови C++, є масштабованим, модульним рішенням. OMNeT++ надає інфраструктуру для конструювання модулів, визначення структури повідомлень та визначення правил їх обробки. Середовище розроблено за принципами вільного розповсюдження і може вільно використовуватися для наукових цілей, а також може використовуватися на платформах Windows і Unix (Linux також). OMNeT++ може використовуватися для проведення дослідження в наступних напрямках:

- моделювання провідних і безпроводних комунікаційних мереж;
- моделювання протоколів роботи мереж;
- моделювання задач теорії черг;
- моделювання багатопроцесорних та інших розподілених комп'ютерних систем;
- тестування архітектури;
- проведення імітаційного моделювання роботи складних програмних систем;

Взагалі, середовище може використовуватися для моделювання будь-яких систем, які допускають дискретне представлення і можуть бути описані у вигляді сутностей, що обмінюються повідомленнями.

Розглянемо узагальнення моделі мережі, яка теоретично досліджувалась у [21, 22].

Модель обробки різнотипних пакетів в графічному вигляді показана на рис. 4 і складається з наступних елементів: користувачі, класифікатор пакетів, модуль обробки пакета для виконання, сховище спільних ресурсів, буфер сервера, сервер, вихідний канал.

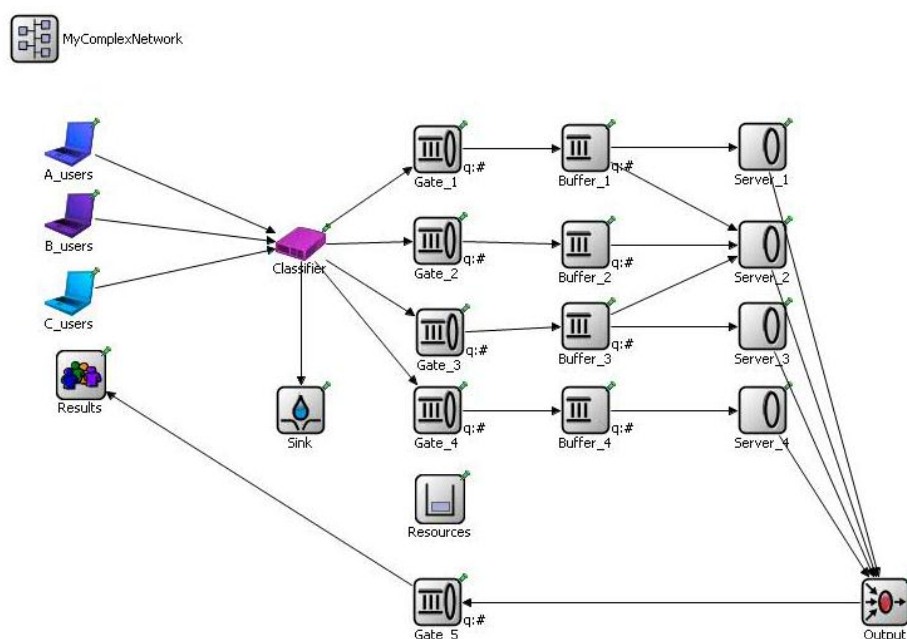


Рис. 4. Модель обробки різнотипних пакетів

Система, що розглядається, призначена для обробки запитів від користувачів (A_users , B_users , C_users на рис. 4). Користувачі генерують пакети з запитом, які мають різні пріоритети (від 0 – найвищий, до 3 – найнижчий). Пакети надходять на вхід класифікатора (Classifier), який проводить аналіз їх пріоритетності. В залежності від конкретного значення пакет буде надісланий на відповідний модуль обробки ($Gate_1$ – $Gate_4$) для підготовки до виконання або відкинутий як невідповідний (Sink). Ця процедура використовує певний спільний ресурс (наприклад, пропускну здатність вхідного каналу або частину пам'яті). Спільний для всіх модулів обробки ресурс показано на рис. 4 (Resources). До виконання запиту на серверах ($Server_1$ – $Server_4$) пакети очікують у відповідному буфері ($Buffer_1$ – $Buffer_4$). Після виконання пакети з результатами потрапляють у модуль обробки ($Gate_5$) та надсилаються користувачам.

Позначимо частоту генерації пакетів як α_i , де індекс – це пріоритет пакета $i \in I$, $I = \{0, 1, 2, 3\}$. Довжину черги в модулі обробки позначимо $q_i(t)$. Тоді процес обробки буде описуватися потоковою моделлю:

$$\dot{q}_i(t) = \alpha_i - u_i(t), \quad i \in I,$$

де керування $u_i(t)$ описує схему черговості обробки пакетів. При цьому на керування накладається обмеження

$\sum_{i=0}^3 u_i(t) \leq \mu$. Аналогічно позначимо затримки у буферах $p_i(t)$ та отримаємо модель

$$\dot{p}_i(t) = u_i(t) - Bv(t), \quad i \in I,$$

де матриця B описує взаємозв'язки між буферами та серверами (які можуть бути, як показано на рис. 3, зв'язувати декілька буферів і один сервер). Для моделі, що розглядається, матриця B має вигляд

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Керування $v = (v_1, \dots, v_4)$, $|v_i| \leq v$ описує обробку пакетів серверами.

Функції обробки пакетів $u(t)$, $v(t)$ можуть задаватися по-різному. Розглянемо один із розповсюджених підходів. Обробка пакетів на вході здійснюється за пріоритетністю, тобто в першу чергу за наявності ресурсів обслуговуються більш пріоритетні пакети. Внаслідок цього пакети найвищого пріоритету обслуговуються майже неперервно, пакети найнижчого – тільки за наявності ресурсу.

Виконаємо моделювання для ситуації коли ресурсів більше, ніж необхідно для обробки всіх пакетів, і при цьому частота генерації пакетів всіх користувачів дорівнює α . Кожен користувач генерує пакети різного пріоритету з рівною ймовірністю. Тому можна оцінити частоту генерації пакетів певного пріоритету одним користувачем як четверту частину від загальної. Отже, пакети кожного пріоритету подаються на вхід системи з частотою $\frac{3\alpha}{4}$.

На рис. 5 – 7 показана залежність довжини черги $q_i(t)$ від часу при різних значеннях μ в таблиці.

Таблиця

№	Значення	Рисунок
1	$\frac{\mu}{4} > \frac{3\alpha}{4}$	5
2	$\frac{\mu}{4} \approx \frac{3\alpha}{4}$	6
3	$\frac{\mu}{4} < \frac{3\alpha}{4}$	7

Перший випадок відповідає ситуації, коли система обробляє пакети швидше, ніж вони встигають накопичуватися у чергах. У цьому разі статистичні коливання не призводять до суттєвого зростання величини черги.

У разі, коли $\frac{\mu}{4} \approx \frac{3\alpha}{4}$, черги не будуть сильно зростати, однак статистичні коливання будуть накопичуватися (рис. 6).

Нарешті, рис. 7 ілюструє ситуацію, коли система не встигає обробити частину пакетів, внаслідок чого черги зростають.

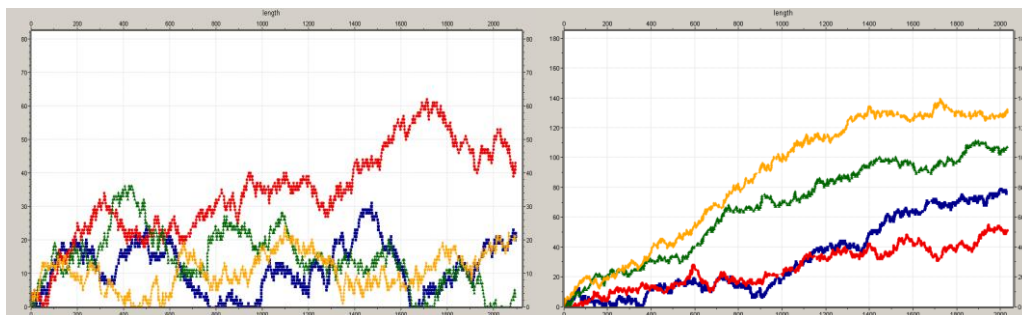


Рис. 5. Стабільна робота системи

Рис. 6. Повільне накопичення помилок

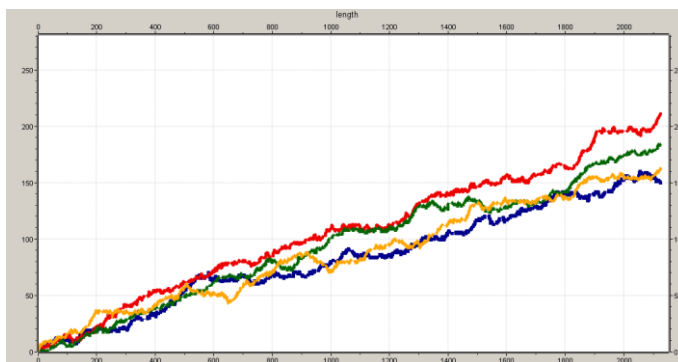


Рис. 7. Зростання черг

Висновки

Для вирішення задач моделювання процесів керування комп'ютерними мережами в умовах конфлікту автором використано поточкову модель та розширено її для врахування конфлікту. Описана поведінки нападника і захисника у термінах диференціальної гри. Проведена формальна постановка й аналіз отриманої гри та знайдені умови на можливості гравців і початковий стан системи, за яких гра може бути закінчена.

1. *Huitema C.* Routing in the Internet. Prentice-Hall, Inc., Englewood Cliffs, N J, 1999.
2. *Willinger W., Paxson V. and other.* Long-range dependence and data network traffic // Theory and applications of long-range dependence. – 2003. – N 4. – P. 373–407.
3. *Altman E., Boulogne T. and other.,* A survey on networking games in telecommunications // Comp. and Oper. Research. – 2006, – **33**, – N 2. – P. 286 – 311.
4. *Kendall D.G.* Some problems in the theory of queues // J. Roy. Statist. Soc. – 1951. – N 13. – P. 151 – 173.
5. *Cohen J. W.* The Single server queue. – Noth. Holland, Amsterdam, 1969 – 657 p.
6. *Koenigsberg E.* Cyclic queues. // Oper. Research Quart. – 1958. – N 9. – P. 22 – 35.
7. *Jackson J.R.* Networks of waiting lines. // Oper. Research. – 1957. – N 5. – P. 518 – 521.
8. *Kleinrock L.* Queuing systems. Wiley Intersciences. – New York, 1975. –**2**, –475 p.
9. *Perros H.G.* Queuing Networks with Blocking. – Oxford University Press, 1994, 563 p.
10. *Kingman J.F.C.* The single server queue in heavy traffic. // Proc. Cambridge. Philos. Soc. – 1961. – N 57. – P. 902 – 904.
11. *Meyn S.P., Tweedie R.L.* Markov Chains and Stochastic Stability. Springer-Verlag, 1993. – 558 p.
12. *Kumar P.R., Seidman T.I.* Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems // IEEE Trans. Automat. Control. – 1990. – N 35, P. 289 – 298.

13. Рыбко А. Н., Столяр А. Л. Об эргодичности случайных процессов, описывающих функционирование открытых сетей массового обслуживания // Пробл. передачи информ. – 1992. – № 28. – С. 3–26.
14. Dai J.G. On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Limit Models // Annals of Applied Probability. – 1995. – N 5. – P. 49 – 77.
15. Harrison, J. M. Brownian models of queueing networks with heterogeneous customer populations // Institute for Mathematics and Its Applications. – 1988, **10**. – 147 p.
16. Bauerle N. Asymptotic optimality of Tracking-policies in stochastic networks // Annals of Applied Probability. – 2000. – N 10. – P. 1065 – 1083.
17. Bauerle N., Rieder U. Optimal control of single-server fluid networks // Queueing Systems. – 2000. – N 35. – P. 185 – 200.
18. Meun S. Control Techniques for Complex Networks. – Cambridge University Press, 2007. – 582 p.
19. Maglaras C. Dynamic scheduling in multiclass queueing networks: stability under discrete-review policies // Queueing Systems. – 1999. – N 31. – P. 171 – 206.
20. Андон П.І., Ігнатенко О.П. Атаки на відмову в мережі Інтернет: опис проблеми та підходів щодо її вирішення. – Київ, 2008. – 50 с. – (Препр./ Ін-т програмних систем НАН України, 2008).
21. Ігнатенко О.П. Конфліктна задача взаємодії двох гравців у відкритому інформаційному середовищі // Проблеми програмування. – 2009. – № 2. – С. 56 – 65.
22. Ігнатенко О.П. Моделювання обслуговування користувачів у мережі з одним сервером за умов конфлікту // Проблеми програмування. – 2009. – № 3. – С. 66 – 72.