

ОПТИМІЗАЦІЯ ФАЙЛООБМІНУ В ПІРІНГОВИХ МЕРЕЖАХ

Петров О.М.

Харківський національний університет радіоелектроніки
61166, Харків, пр.Науки, каф.Інфокомунікаційної інженерії, тел. (057) 702-13-20,
E-mail:: oleksandr.petrov1@nure.ua

The peer network is based on the equality of participants and is characterized by the fact that its elements can be interconnected. One area of application of data network technology is file sharing. The task of optimizing file sharing can be considered as a scheduling problem. It examines the tasks in which it is necessary to order or, in other words, to determine the sequence of performance of a set of works.

Вступ

З розвитком технологій необхідність мати доступ до величезної кількості інформації, що лежить на інших комп'ютерах, спочатку породила локальні мережі. Це допомогло на деякий час, але об'єднати таким чином велику кількість комп'ютерів виявилось неможливо. Можливість же доступу до файлів, можливо наявних у кого-то ще, була дуже приваблива. Тому була запропонована система, що дозволяє приймати і передавати файли з кожного комп'ютера, з яким є зв'язок і на якому є певне програмне забезпечення. Такі системи отримали назву P2P або пірінгових (peer-to-peer) мереж.

Пірінгова мережа будується на основі інтернет-з'єднання, використовуючи особисті протоколи передачі даних. P2P технології покращують масштабність системи при низьких затратах реалізації. На відміну від традиційної клієнт-серверної архітектури, піри в мережі виступають в якості клієнта (ліч-скачує) і сервера (сід-роздає). Піри не тільки скачують файл з мережі, а й роздають завантажений файл іншим користувачам в мережі. Частина файлів обмінюються через прямі з'єднання між пірами. Іншою важливою властивістю пірінгових мереж є – можливість упорядкованості фрагментів, переривання, входжень і виходів з мережі в будь-який момент часу.

Так як в даний час Інтернет отримує все більшу і більшу поширеність, використання пірінгових мереж є більш актуальним з кожним днем.

В праці розв'язується задача оптимізації файлообміну в пірінгових мережах.

Основні терміни, вихідні величини та рішення задачі теорії розкладу

Задачу оптимізації файлообміну, можна розглядати як задачу теорії розкладів. Вона - є частиною дослідження операцій та досліджує задачі, в яких необхідно впорядкувати або, іншими словами, визначити послідовність виконання сукупності робіт.

Основним поняттям теорії розкладів - є поняття операції. Операцію можна розглядати як елементарну задачу, що підлягає виконанню. Кожна операція характеризується:

- 1) індексом приналежності до певної роботи;
- 2) індексом приналежності до певної машини;
- 3) числом, що представляє собою тривалість операції.

По першому індексу всі множини операцій розбиваються на систему непересічних підмножин, які називаються роботами. Розбиття вихідної множини по-другому індексу призводить до взаємно непересічних підмножин операцій, що відносяться до певних машин.

Для кожної роботи задається послідовність складових її операцій (що визначається технологічним процесом). Таке часткове впорядкування операцій здійснюється завданням відносини порядку. Якщо операція X повинна бути здійснена раніше, ніж Y, то кажуть, що X передре Y. Це записується у вигляді: $X < Y$ або $Y > X$. Співвідношення порядку транзитивне, якщо $X < Y$, $Y < Z$, то $X < Z$. Будемо вважати, що операція X безпосередньо передре операції Y і записувати це у вигляді $X << Y$ або $Y >> X$, якщо $X < Y$ і немає операції Z, такої, що $X < Z < Y$.

Часто буває зручно представляти згадані співвідношення у вигляді орієнтованого

графа. Вершины (вузли) графа зображають операції, а дуги - відношення безпосереднього передування. Дві вершини пов'язані відношенням порядку, якщо існує шлях між ними.

Машиною будемо називати пристрій, здатний виконати все, що пов'язано з деякою операцією, системою обслуговування - множина всіх машин, які використовуються для виконання деякої множини операцій. Сукупність машин, робіт (операцій) і дисциплін призначення операцій відповідним машинам називається процесом обслуговування.

Складання розкладу для процесу обслуговування означає, що для кожної операції на тимчасовій осі задається ділянка, коли ця операція повинна виконуватися відповідної машиною.

Тобто складання розкладу може розглядатися як завдання впорядкування операцій, які виконуються кожною машиною.

Розглянемо величини, які використовуються в подальшому як критерії оцінки розкладів, та співвідношення між ними.

Спочатку розділимо вихідні та шукані величини задачі. Перші визначаються специфікою розв'язуваної задачі, інші є результатом складання розкладу. Щоб підкреслити цю різницю, під час визначення цих величин рядкові латинські літери, а при позначенні шуканих - прописні. Наприклад: x, y, z та X, Y, Z .

Постановка задачі в теорії розкладів починається з опису системи машин і множини робіт. Для найпростішого процесу обслуговування система машин повністю описується їх числом. Нехай система складається із m машин, занумеруємо їх $1, 2, \dots, m$. Занумеруємо також роботи числами від 1 до n , і нехай:

r_i - момент готовності (момент появи або момент надходження) роботи. Ця величина являє собою момент надходження i -ї роботи в систему з деякого зовнішнього джерела.

r_i є мінімально можливий час початку першої з операцій роботи i , $i=1, \dots, n$;

d_i - плановий (директивний) строк. Ця величина являє собою момент, до якого i -та робота повинна бути виконана. Іншими словами, d_i являє собою директивний час закінчення операції, заданий деякими зовнішніми по відношенню до даної системи причинами.

Допустима тривалість проходження роботи в системі дорівнює $a_i = d_i + r_i$. Досить задати дві з трьох величин r_i, d_i, a_i . Теоретично байдуже, які дві з них задані, а яка знаходиться з наведеного співвідношення, тому для кожної роботи всі вони вважаються заданими.

Задача теорії розкладу вважається заданою, якщо визначені:

- 1) належні виконанню роботи та операції;
- 2) кількість та типи машин, виконуючих операцію;
- 3) порядок проходження машин;
- 4) критерії оцінки розкладів.

Загальна задача теорії розкладу записується так: $n/m/G/F_{\max}$ - впорядкувати n робіт у довільній системі із m машин так, щоб мінімізувати максимальну тривалість проходження роботи.

Пошук оптимального або близького до оптимального розкладу здійснюється за допомогою одного з 4-х підходів (математичного програмування; комбінаторного; евристичного; статичного (імовірнісного)).

У роботі був обраний перший підхід - математичного програмування, в термінах лінійного цілочисельного і динамічного програмування для 3х пірів та чотирьох програм. Була сформульована задача лінійного цілочисельного програмування.

Цільова функція - це середньозважений час початку зчитування фрагментів. Кожна машина в один момент часу може виконувати не більше однієї роботи, з цього випливає, що для кожної пари робіт виконується лише одна з нерівностей:

- 1) Виконання робіт j передуватиме виконанню робіт i .

2) Виконання робіт і передуює виконанню робіт j .

Таке обмеження типу «або - або» не можна описати в рамках звичайного лінійного програмування і вимагає введення цілочисельних змінних. Введемо цілочисельну змінну Y_{ij} , яка може приймати значення 1 або 0. Якщо Y приймає значення 0, виконується перша умова, якщо 1 то друга.

Також в роботі вирішена задача динамічного програмування, яку можна інтерпретувати як задачу про найкоротший шлях на графі, де довжини дуги - це тривалість виконання роботи.

Для задачі про найкоротший шлях використовується рекурсивне рівняння Беллмана:

$$f_n(s) = \min[c_{sj} + f_{n-1}(j)], n = 1, \bar{k} \quad (1)$$

де k - число етапів на шляху початкової вершини в кінцеву;

c_{sj} - вартість по дузі (S, j) ;

$f_n(s)$ - вартість, що відповідає стратегії мінімальних завмирань для шляху з пункту s , якщо до кінцевого пункту залишилося n етапів;

$j_n(s)$ - рішення, що дозволяє досягти;

f - значення цільової функції;

S - значення залежить від стану системи, індекс n вказує, скільки етапів залишається до кінця шляху.

На основі вихідних даних: матриці розподілу програм по пірамі та матриці тривалостей зчитування фрагментів будується мережевий граф, де кожна вершина описує процес виконання конкретного фрагменту на конкретному пірі, а довжини дуг визначають тривалість такого зчитування. Був сформований граф виконання робіт, а також граф кожного з етапів виконання роботи.

Рішення задачі динамічного програмування виконується в кілька етапів. Використовуючи оптимальне рішення на кожному етапі, знаходимо найкоротші шляхи, як послідовність вершин графу. Отримані послідовності вершин являють собою порядок скачування фрагментів програм на кожній із машин.

Висновки

Вирішена задача планування файлообміну в пірінговій мережі з трьома пірамі та чотирма програмами з використанням методів лінійного цілочисельного та динамічного програмування. Задача лінійного цілочисельного програмування потребує набагато більшого часу на етапі підготовки, аніж задача динамічного програмування, але її реалізація набагато легше.

З результатів моделювання випливає, що як середньозважений час зчитування файлів, так і умовний середній оптимальний час зчитування файлів практично не залежить від закону розподілу значень вагових коефіцієнтів цільової функції і від закону розподілу значень тривалостей зчитування фрагментів файлів пірамі і залишається постійним. Цей результат має практичне значення в тому, що на практиці не потрібно аналізувати невідомий закон розподілу випадкових тривалостей обробки фрагментів файлів пірамі на підставі статистики, достатньо знати середнє вибіркоче. Крім того, показано, що моделі та методи теорії розкладу – адекватні при рішенні задач в області поліпшення якісних показників пірінгових мереж.

Як показав імітаційний аналіз використання оптимізації дозволило зменшити час зчитування файлів в 2 рази.

Список літератури:

1) Гордон В. С. Задачи теории расписаний с заданным частичным порядком / В. С. Гордон, А. Б. Долгий. // Автоматика и телемеханика. – 2010. – №10. – С. 15–27.

Четвертая международная научно-техническая конференция
Проблемы электромагнитной совместимости перспективных беспроводных сетей
связи

- 2) Финкова М. В. Пиринговые сети / М. В. Финкова. – СПб: Наука и техника, 2006. – 277 с.
- 3) Поповский В.В. Математические основы управления и адаптации в телекоммуникационных системах: учебник / В. В. Поповский, В. Ф. Олейник. – Харьков: СМІТ, 2011. – 362 с.
- 4) Математичні основи теорії телекомунікаційних систем: навч. Посібник / [В. В. Поповський, С. О. Сабурова, В. Ф. Олійник та ін.]; під заг. ред. В.В. Поповського. – Харків: СМІТ, 2006. – 564 с.