

СИНТЕЗ ОПТИМАЛЬНОЇ СТРУКТУРИ ДОКУМЕНТООБІГУ ОРГАНІЗАЦІЙ

д.т.н. Т.В. Бутько, Т.Г. Білова

Запропонована методика оптимізації структури електронного документообігу організацій різних форм власності, яка базується на засобах теорії масового обслуговування, статистичного моделювання та теорії графів.

Правильна організація документообігу сприяє оперативному проходженню документів в апараті управління, рівномірному завантаженню підрозділів та посадових осіб, оказує позитивний вплив на процес управління взагалі. Результатом того, що в організації не налагоджена чітка робота з документами, може бути зниження ефективності управління, тому що воно залежить від якості та достовірності, оперативності прийому - передачі інформації, правильного ставлення довідково - інформаційної служби, чіткої організації пошуку, зберігання та використання документів.

Документообіг в організації здійснюється у вигляді потоків документів, які циркулюють між пунктами обробки інформації. У сучасній організації такими пунктами є комп'ютери, встановлені на робочих місцях персоналу та з'єднані у локальну комп'ютерну мережу. Якщо ця мережа охоплює всі місця діловодчого персоналу, то з'являється можливість використовувати мережу для переміщення документів. У такому випадку систему документообігу можна уявити, як інформаційну мережу з вхідним потоком інформації, потоком обслуговування у мережі та вихідним потоком документів (рис. 1).

Для оптимізації існуючої системи документообігу дослідимо динаміку інформаційної мережі, визначення структури і параметрів інформаційних потоків. Критерієм оптимізації можна вибрати наступну цільову функцію:

$$C = t_{03}C_{пот} + \sum_i t_{прi}^{kc} C_i^{kc} + \sum_j t_{прj}^n C_j^n \rightarrow \min, \quad (1)$$

де t_{03} - середній час очікування документу на вході системи;

$C_{пот}$ - витрати від очікування документу;

$t_{прi}^{kc}$ - час простою каналів зв'язку;

C_i^{kc} - витрати на експлуатацію каналів зв'язку в одиницю часу;

t_{npj}^n - час простою обслуговуючого персоналу;

C_j^n - витрати на заробітну плату обслуговуючого персоналу.

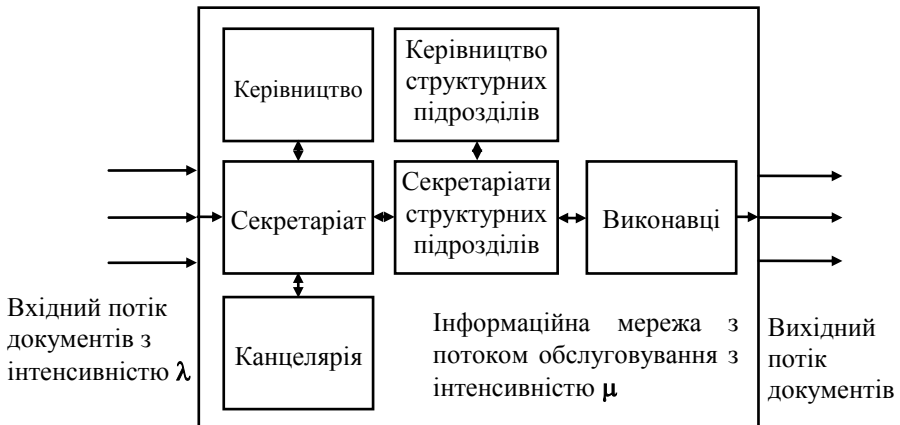


Рис. 1. Система документообігу як інформаційна мережа

Розглянемо два види інформаційних потоків: вхідний інформаційний потік та потік обслуговування у мережі.

За статистичними даними аналізу вхідного потоку інформації визначається вибірка $\{t_i\}$, $i = \overline{1, n}$, часу між двома послідовними надходженнями заявок на вхід мережі. При зборі статистичних даних повинна дотримуватися вимога до репрезентативності вибірки.

В результаті попередніх спостережень було виявлено, що вхідний потік заявок, як правило, є потоком Ерланга k -го порядку та його можна описати розподілом Ерланга

$$f(t) = \frac{(\lambda k)(\lambda k t)^{k-1}}{(k-1)!} e^{-\lambda k t}, \quad (2)$$

де k - додатне ціле число; λ - інтенсивність потоку.

Потоки Ерланга створюють цілу гаму потоків з різним ступенем упорядкованості, від найпростішого ($k = 1$) до регулярного ($k \rightarrow \infty$). Особливим випадком потоку Ерланга є найпростіший потік (тобто ординарний, стаціонарний та ергодичний), який можна описати законом Пуассона

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (3)$$

де $P_k(t)$ - ймовірність того, що на довільно обраній ділянці часу тривалістю Δt надійде k вимог.

У найпростішому потоці інтервал часу між будь - якими двома сусідніми випадками розподілено за експотенційним законом з параметром λ

$$f(t) = \lambda e^{-\lambda t} . \quad (4)$$

Для визначення типу розподілу вхідного потоку заявок амплітуда коливань проміжків часу розбивається на N інтервалів

$$(N = 1 + [1.33 \lg n]) \quad (5)$$

та визначається m_i - кількість попадань випадкової величини T часу між сусідніми заявками в i - й інтервал. Будується гістограма, яка описує емпіричний розподіл інтервалів часу між надходженнями заявок. За виглядом гістограми висувається гіпотеза про вид розподілу. Для підбору кривої, що апроксимує розподіл, λ розраховується як величина, зворотна математичному очікуванню інтервалів між сусідніми заявками \bar{T} , тобто

$$\lambda = 1/\bar{T} . \quad (6)$$

Для доказу достовірності припущення про вигляд розподілу величини T треба перевірити близькість емпіричного та теоретичного розподілів частотей за критерієм Пірсона χ^2 :

$$\chi^2 = \sum_{i=1}^N \frac{(m_i - m_{ip})^2}{m_{ip}} , \quad (7)$$

де m_{ip} - розрахункові (теоретичні) частоти повторення проміжків у кожному i - му інтервалі, для експоненційного розподілу визначаються

$$m_{ip} = n\lambda \Delta t e^{-\lambda \bar{t}_i} , \quad (8)$$

де Δt - довжина інтервалу групування випадкової величини;

\bar{t}_i - середина i - го інтервалу групування.

За правилом Романовського оцінка близькості емпіричного розподілу до теоретичного має вигляд

$$\left| (\chi^2 - \omega) / \sqrt{2\omega} \right| < 3 , \quad (9)$$

де ω - число ступенів свободи, $\omega = N - 2$.

Якщо нерівність виконується, то гіпотеза про вигляд розподілу приймається, інакше висовується інша гіпотеза.

Потік інформації у мережі досліджується на основі вибірки часу обслуговування заявки. Час обслуговування - випадкова величина із законом розподілу

$$F(t) = P(T < t) , \quad (10)$$

де $P(T < t)$ - ймовірність того, що час обслуговування не перебільшує деякої заданої величини t .

Основною характеристикою часу обслуговування є інтенсивність обслуговування μ , величина, зворотна середньому часу обслуговування. Для визначення закону розподілу часу обслуговування будується гістограма

частостей. За виглядом гістограми висувається гіпотеза про вигляд розподілу часу обслуговування. Справедливість гіпотези перевіряється з допомогою критерію χ^2 за формулами (7) - (9).

Якщо процес, що протікає у мережі, є марківським (тобто його характеристики в майбутньому залежать тільки від стану його в теперішній момент часу та не залежать від того, як він протікав у минулому), то математичний аналіз системи значно полегшується. Тоді є можливість описати роботу системи за допомогою апарату диференціальних (а у випадку сталого режиму - лінійних алгебраїчних) рівнянь та виразити основні характеристики системи через параметри потоку заявок та потоку обслуговування. Коли розподіл вхідного потоку є розподілом Ерланга k - го порядку, треба використовувати метод псевдостанів. У випадках, коли вхідний потік та потік обслуговування не підпорядкований розподілу Ерланга, доцільно використовувати метод статистичного моделювання (метод Монте - Карло) [2].

У зв'язку із складністю систем документообігу найбільш ефективною схемою дослідження є поєднання аналітичних та статистичних методів: за допомогою аналітичних методів будуються моделі окремих підсистем, а потім з них синтезується статистична модель системи.

На основі аналізу існуючих систем документообігу, якій було проведено в організаціях різних форм власності пропонується наступна класифікація топології інформаційних мереж: шина (рис. 2а); зірка (рис. 2б); кільце (рис. 2в); дерево (рис. 2г); змішана (рис. 2д).

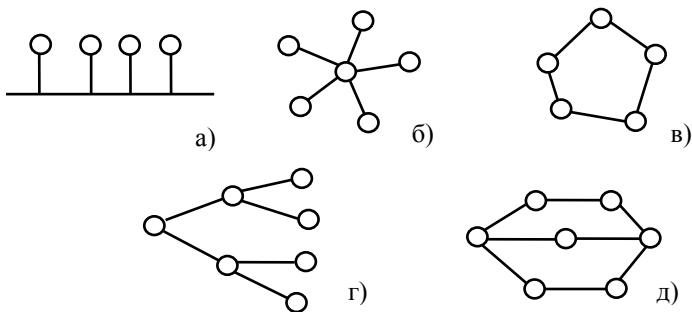


Рис.2. Типові топологічні схеми організації документообігу

Застосування цієї методики дослідження і оптимізації документообігу доцільно розглянути на прикладі системи обробки запитів виборців у Миську Раду. Вхідний потік інформації - це запити громадян, що надходять у письмовій формі або по телефону. Потік обслуговування - обробка запитів у відділі з роботи із зверненнями громадян та передача їх у підрозділи виконавчої влади, які відповідають за вирішення конкретних питань.

За попереднім аналізом вхідного потоку заяв та часу обслуговування були зроблені такі висновки:

1) потік заявок, який надходить на вхід системи, є псевдостационарним, він змінюється в залежності від часу року. У зв'язку з цим стає питання про визначення моментів часу, на яких потік є стационарним;

2) потік заявок на стационарних інтервалах часу є найпростішим з інтенсивністю $\lambda=3 \text{ г}^{-1}$ (гістограма розподілу наведена на рис. 3), а потік обслуговування - потоком Ерланга 2 - го порядку з інтенсивністю обслуговування $\mu=0.025 \text{ г}^{-1}$.

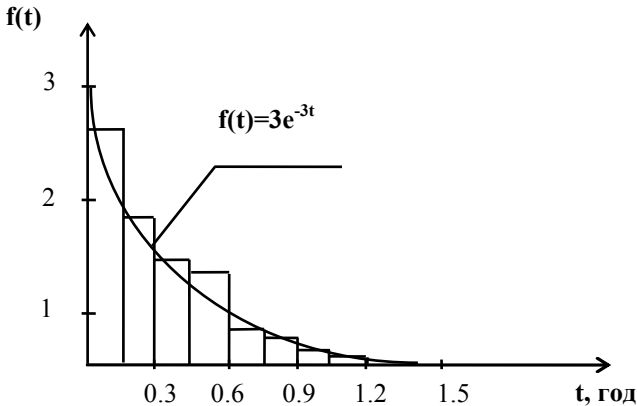


Рис.3. Гістограма розподілу інтервалів часу між надходженнями заявок

Аналіз роботи системи показав, що:

- 1) обслуговування заявок чиниться з абсолютними пріоритетами;
- 2) вхідний потік заявок є необмеженим;
- 3) система обробки заявок є багатофазною системою з очікуванням.

Важливою є задача приведення в кількісну відповідність потоку заявок з кількісними можливостями (потужністю) інформаційної мережі. За попереднім аналізом структура документообігу системи є змішаною (рис. 1д).

Відобразимо систему обробки запитів виборців у вигляді граф - моделі (мережі) $G = (I, U)$, де I - вершини графу (пункти обробки інформації), U - ребра графу (потоки інформації між пунктами). Ребра графу характеризуються функцією пропускну здатності каналів зв'язку, яка ставить у відповідність дузі $(i, j) \in U$ графу G пропускну здатність $r(i, j)$ каналів зв'язку.

Функція потоку у мережі з одним джерелом s та одним стоком t співставляє з кожною дугою (i, j) ціле число $x(i, j)$ та має властивості:

$$0 \leq x(i, j) \leq r(i, j), \quad (i, j) \in U; \quad (11)$$

$$\sum_i x(i, k) - \sum_j x(k, j) = 0, k \neq s, t, k \in I; \quad (12)$$

$$\sum_k x(s, k) = \sum_k x(k, t) = v. \quad (13)$$

Потік $x(i, j)$ вздовж дуги (i, j) - кількість інформації, яка проходить через канал в одиницю часу. Потік в мережі - сукупність $\{x(i, j)\}$ потоків по усім дугам мережі.

Основною задачею при визначенні потужності мережі є визначення величини найбільшого потоку, допустимого у мережі при даній функції пропускної здатності. За теоремою Форда - Фалкерсона, для будь - якої мережі максимальна величина потоку із s в t дорівнює пропускної здатності мінімального розрізу, який відділяє s від t , тобто задача зводиться до визначення мінімального розрізу мережі

$$r(R, \bar{R}) = \sum_{i \in R, j \in \bar{R}} r(i, j), \quad (14)$$

де R та \bar{R} - підмножини усіх вершин мережі G , що не перекреслюються.

Для практичних розрахунків найбільш зручно використовувати для визначення мінімального розрізу алгоритм Форда [1].

З урахуванням приведеної методики зараз проводяться роботи по оптимізації системи електронного документообігу в організаціях різних форм власності.

ЛІТЕРАТУРА

1. Балашевич В.А. Математические методы в управлении производством. - Минск: Вышэйшая школа, 1976. - 336 с.
2. Вентцель Е.С. Исследование операций: задачи, принципы, методология. - М.: Наука, 1980. - 208 с.
3. Черняк Ю.И. Цели и методы исследования потоков информации // Исследование потоков экономической информации. - М.: Наука, 1968. - С. 7 - 28.