

## НЕКОТОРЫЕ ВОПРОСЫ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ СЕТЕВЫХ ГРАФИКОВ

Ю. Н. ЕФИМОВ

(Представлена научным семинаром вычислительной лаборатории)

Эвристический алгоритм оптимизации расписания работ сетевого графика в системе СПУ с ограниченными ресурсами [1] приводит в большинстве случаев к практически удовлетворительным результатам. Однако могут встретиться сетевые графики, для которых алгоритм оптимизации не дает никакого эффекта. Для таких сетевых графиков описанный в [1] коэффициент улучшения близок к единице. Естественно, что в подобных случаях нет смысла производить сравнительно трудоемкую процедуру оптимизации. Поэтому желательно было бы заранее оценить, хотя бы с вероятностной точки зрения, величину коэффициента улучшения  $K_{\text{л}}$ . Вопросам предварительной оценки коэффициента улучшения и посвящена эта статья.

В общем случае  $K_{\text{ул}}$  зависит от топологии сети, длительности работ  $t_{ij}$  и интенсивности потребления ресурса на работах  $v_{ij}$ . Например, можно подобрать такие величины  $t_{ij}$ , чтобы все работы сетевого графика были «критическими» или подобрать такие значения  $v_{ij}$ , чтобы график загрузки оборудования был равномерным. Понятно, что в этих случаях нет смысла говорить об оптимизации работ.  $K_{\text{ул}}$  зависит также от соотношения числа событий  $M$  и работ  $N$  сетевого графика. В самом деле, при  $N = M - 1$  сетевой график вырождается в одну цепь из «критических» работ. С ростом числа  $N$  при  $M = \text{const}$ , т. е. с увеличением коэффициента сложности сетевого графика ( $K_{\text{сл}} = \frac{N}{M}$ ), появляются разветвления, а с ними и возможность оптимизации. Причем априори можно утверждать, что с ростом  $K_{\text{сл}}$  значение  $K_{\text{ул}}$  будет тенденциозно убывать. При максимально возможном числе работ  $N_{\text{max}} = \frac{M(M-1)}{2}$  сетевой график представляет собой полный граф [2] и достигается максимальный эффект применения оптимизации. Попробуем оценить математическое ожидание коэффициента улучшения ( $M[K_{\text{ул}}]$ ) в зависимости от  $K_{\text{сл}}$ .

Прежде всего нас будет интересовать закон распределения  $K_{\text{ул}}$ . Используя обозначения [1], имеем

$$K_{\text{ул}} = \frac{\max R_n(t)}{\max R_0(t)},$$

$$\max R_n(t) = \sum_{(ij) \in F_n(t)} v_{ij};$$

$$\max R_0(t) = \sum_{(ij) \in F_0(t)} v_{ij};$$

$F_0(t)$  и  $F_n(t)$  — фронты работ (до и после оптимизации соответственно), на которых потребляется максимальное количество ресурса.

Произведя замену,

$$\begin{aligned} \max R_n(t) &= X, \\ \max R_0(t) &= Y, \\ K_{ул} &= Z, \end{aligned}$$

получим

$$Z = \frac{X}{Y}. \tag{1}$$

Интенсивности потребления ресурса работами, принадлежащими определенному фронту, являются величинами случайными, следовательно,  $X$  и  $Y$  представляют собой сумму случайного числа случайных величин. Так как суммируемые величины имеют одинаковую природу, можно применить теорему Ляпунова [3], согласно которой случайные величины  $X$  и  $Y$  распределены по нормальному закону с математическими ожиданиями  $m_x, m_y$  и дисперсиями  $D_x, D_y$ . Следовательно, функции плотности распределения  $f(x)$  и  $f(y)$  можно считать известными. Плотность вероятности функции ( $Z$ ) двух случайных аргументов определяем по формуле [3].

$$f_z(z) = \int_{-\infty}^{\infty} f(\psi(z,y), y) \left| \frac{\partial \psi(z,y)}{\partial z} \right| dy,$$

где  $\psi(z,y) = X = XY$  — решение уравнения (1). Так как величины  $X$  и  $Y$  сильно коррелированы, то получение аналитического выражения закона распределения системы случайных величин  $(X, Y)$ , а следовательно, и  $f_z(z)$  представляет очень сложную задачу. Поэтому для нахождения закона распределения случайной величины  $K_{ул}$  воспользуемся методом статистических испытаний.

Зададимся сетевым графиком  $G = (I, U)$ , имеющим  $M$  событий и  $N$  работ. Вариантом  $V$  графика  $G$  назовем сетевой график, у которого длительности работ  $t_{ij}$  и интенсивности потребления ресурса  $v_{ij}$  имеют конкретные значения, полученные случайным образом. Случайные числа  $t_{ij}$  и  $v_{ij}$  в интервале  $(\alpha, \beta)$  удобнее всего получить с помощью программных датчиков случайных чисел. Эти датчики основаны на получении последовательности равномерно распределенных чисел в интервале  $(0,1)$  по рекуррентным формулам. Получаемые случайные числа  $\delta_i$  сводятся к интересующему нас интервалу согласно соотношению

$$\gamma_i = \alpha + \delta_i (\beta - \alpha).$$

Используя соотношение

$$\gamma_i = \sigma \left( \sum_{j=1}^{12} \delta_j - 6 \right) + m,$$

можно сформировать последовательность чисел  $\gamma$ , распределенных по нормальному закону с параметрами  $m$  и  $\sigma$ . Получаемые числа  $\gamma_i$  фактически являются псевдослучайными. Но так как свойства последовательности псевдослучайных чисел соответствуют статистическим свойствам случайных величин, распределенных по заданному закону, эти числа вполне могут быть использованы для статистических испытаний.

Исходную информацию о варианте  $V$  представим в виде списка работ ( $ij$   $t$   $U$ ), их длин и интенсивностей потребления ресурса на этих работах.

$$ijt_{ij}\nu_{ij}. \quad (2)$$

Произведем временной анализ сетевого графика, используя итерационный метод [4],

$$t_p(j) = \max[t_p(j), t_p(i) + t_{ij}],$$

$$t_n(i) = \min[t_n(i), t_n(j) - t_{ij}],$$

где  $t_p$  и  $t_n$  — соответственно наиболее ранний и поздний сроки свершения событий.

Находим резервы ( $P$ ) работ по формуле [1]

$$P = t_p(j) - t_p(i) - t_{ij}$$

и формируем информацию в виде, удобном для оптимизации

$$t_n P N t_{ij} \nu_{ij}, \quad (3)$$

где

$t_n$  — время начала работы в исходном расписании  $t_n = t_p(i)$ .

$N$  — одноиндексный номер работы.

Затем проводим процедуру оптимизации и получаем величину  $K_{ул}$ .

Рассчитывая значение коэффициента улучшения для различных вариантов  $V_i$  ( $i=1, 2, \dots, n$ ) графа  $G$ , можно построить гистограмму частот величины  $K_{ул}$ , а затем аппроксимировать ее соответствующим законом распределения. Производя подобные расчеты для сетевых графиков различной сложности, можно определить зависимость закона распределения величины  $K_{ул}$  от значения коэффициента сложности.

По описанной методике были проведены расчеты сетевых графиков на ЭЦВМ «Минск-1». В программе, реализующей процесс подготовки статистического материала (рис. 1), использовался датчик случайных чисел с равномерным распределением в интервале (1,7) для  $t_{ij}$  и (0,7) для  $\nu_{ij}$ . Выбор интервалов обусловлен ограниченностью разрядной сетки машины и принятым допущением об отсутствии фиктивных работ. Иначе необходимо было бы вводить дополнительные условия на случайные величины, формируемые датчиками. Получаемые коэффициенты сложности выводились на перфоратор для дальнейшей статистической обработки. По результатам статистической обработки строились гистограммы частот  $K_{ул}$ , вид которых показан на рис. 2. Как видно из рис. 2,

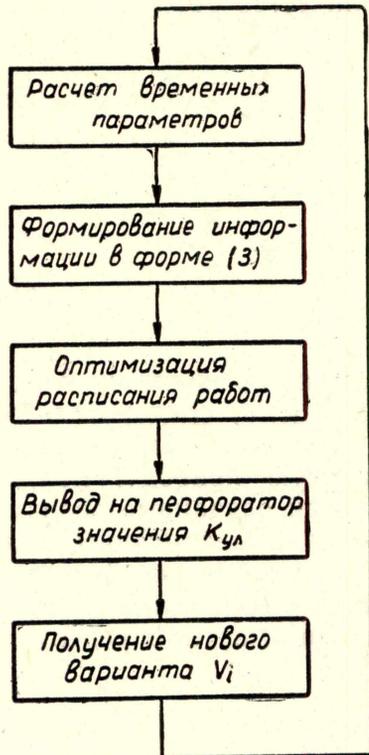


Рис. 1. Блок-схема программы подготовки статистического материала

наиболее соответствующим законом для аппроксимации гистограммы является нормальный закон. Соответствие гипотезы о гауссовом распределении с опытными данными оценивалось по критерию  $\chi^2$ . Для этого по статистическим данным находились математическое ожидание ( $M[K_{ул}]$ ), дисперсия ( $D[K_{ул}]$ ) и среднее квадратическое отклонение  $\sigma$  величины  $K_{ул}$  по формулам [3]:

$$M[K_{ул}] = \sum_{i=1}^n \frac{K_{ул}^i}{n};$$

$$D[K_{ул}] = \frac{\sum_{i=1}^n (K_{ул}^i - M[K_{ул}])^2}{n-1};$$

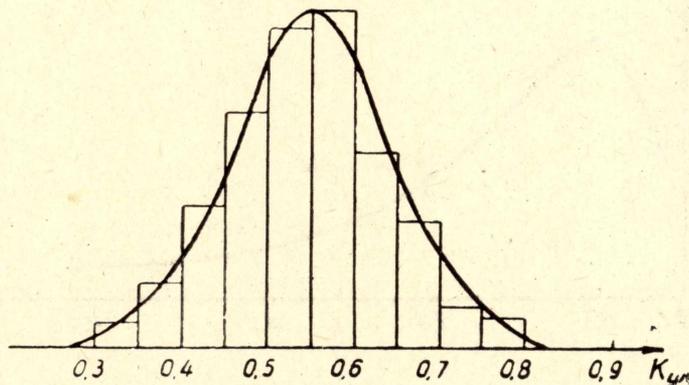


Рис. 2. Пример распределения величины  $K$ :  
 а) гистограмма частот величины  $K$   
 б) график распределения вероятностей идеального гауссова процесса, имеющего такие же параметры.

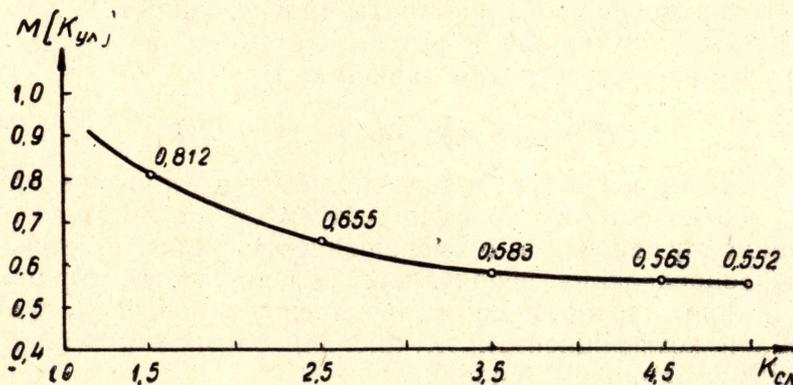


Рис. 3. График зависимости  $M[K_{ул}] = f(K_{сл})$   
 $\sigma = \sqrt{D[K_{ул}]}$ .

Мера расхождения  $\chi^2$  вычислялась по формуле

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i},$$

где

- $m_i$  — количество значений, приходящееся на  $i$ -й разряд,
- $k$  — количество разрядов,
- $p_i$  — вероятность попадания случайной величины  $K_{ул}$  в  $i$ -й разряд, которая находится по формуле

$$p_i = \frac{1}{2} \left[ \Phi \left( \frac{K_{ул}^{i+1} - M[K_{ул}]}{\sigma \sqrt{2}} \right) - \Phi \left( \frac{K_{ул}^i - M[K_{ул}]}{\sigma \sqrt{2}} \right) \right],$$

а  $\Phi(x)$  — известная функция Лапласа.

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

По величине  $\chi^2$  и  $r = k - s$  (при аппроксимации нормальным законом  $s = 3$ ) по таблицам [3] находим вероятность ( $p$ ) того, что величина, имеющая распределение  $\chi^2$  с  $r$  степенями свободы, превзойдет полу-

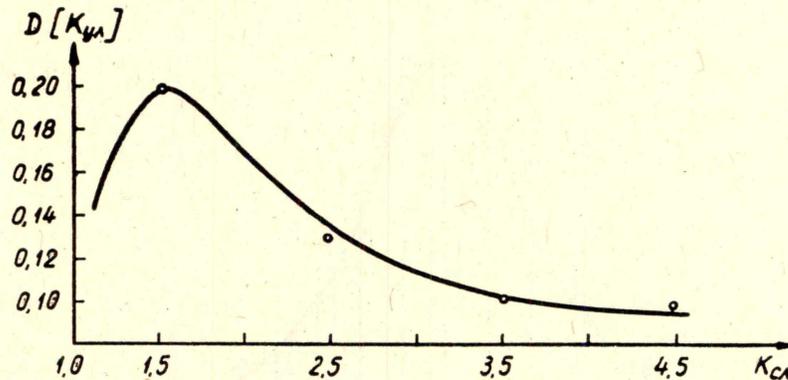


Рис.4. График зависимости  $D[K_{ул}] = f(K_{сл})$ .

ченное значение  $\chi^2$ . Получающиеся вероятности  $p = (0,5 \div 0,8)$  являются достаточно большими и свидетельствуют о том, что гипотеза о гауссовом распределении  $K_{ул}$  не противоречит данным.

По данным, полученным в результате расчетов сетевых графиков различной сложности, построены зависимости

$$M[K_{ул}] = f(K_{сл}), D[K_{ул}] = f(K_{сл})$$

(рис. 3 и 4), которые могут быть использованы для предварительной оценки  $K_{ул}$  по известному коэффициенту сложности сетевого графика.

Как видно из рис. 3 и 4, с уменьшением сложности сетевого графика кривая распределения вероятностей смещается в сторону увеличения  $M[K_{ул}]$  при одновременном увеличении дисперсии. Уменьшение дисперсии при коэффициентах улучшения, близких к единице, объясняется, видимо, влиянием условия:

$$K_{ул} \leq 1.$$

#### ЛИТЕРАТУРА

1. Ю. Н. Ефимов. Оптимизация плана распределения работ в системе СПУ с ограниченными ресурсами. (Настоящий сборник).
2. К. Берг. Теория графов и ее применения. М., ИЛ, 1962.
3. Е. С. Вентцель. Теория вероятностей. М., 1958.
4. Ю. Н. Ефимов. Анализ сетевых графиков на ЭЦВМ «Минск-1». Научный отчет № Р-767, 1967.