2.2. Минимальное общее время выполнения неоднородных конкурирующих процессов в асинхронном режиме

Обозначим минимальное общее время выполнения n, $n \geq 2$, неоднородных процессов с момента начала выполнения первого и до момента завершения последнего в асинхронном режиме, который определяется условиями 1–5, через $T_{сн}^{ас}(p,n,s)$. Предполагается, что каждый из процессов является сосредоточенным.

Рассмотрим следующие случаи.

1) $n \leq p$, т.е. число конкурирующих процессов не превышает числа процессоров многопроцессорной системы.

Тогда, если $n = p$, то достаточно рассмотреть выполнение p процессов на p процессорах. Матрица T^c будет иметь вид $T^c = [t_{ij}]$, где $i = \overline{1,p}$, $j = \overline{1,s}$. Заметим, что при $n < p$ достаточно взять только n процессоров, а остальные $p - n$ не будут задействованы, а матрица $T^c = [t_{ij}]_{n \times s}$.

По матрице T^c строим сетевой вершинно–взвешенный граф G^c с числом вершин $p \times s$ следующим образом: вершинам и дугам графа поставим в соответствие узлы и дуги прямоугольной решетки $p \times s$. Нумерацию вершин и их веса определим соответствующими элементами матрицы $[t_{ij}]$, $i = \overline{1, p}$, $j = \overline{1, s}$, причем начальная вершина будет иметь вес t_{11}, а заключительная – t_{ps}. Горизонтальные дуги в графе G^c означают линейный порядок выполнения блоков, а вертикальные – линейный порядок выполнения процессов (рис. 2.1).

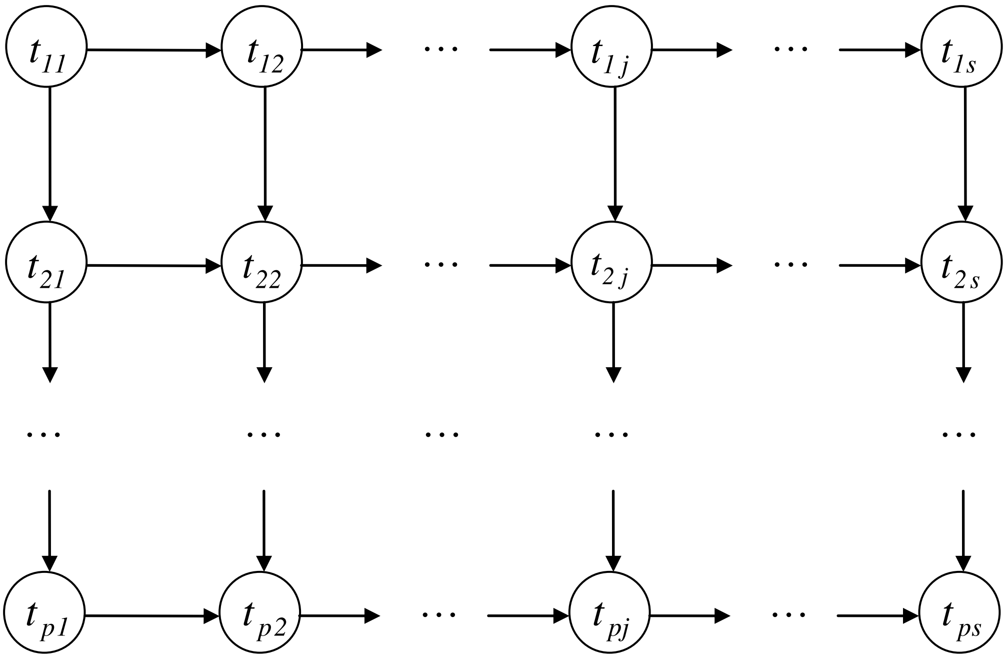

Рис. 2.1 – Сетевой вершинно–взвешенный граф G^c

Имеет место следующая теорема.

Теорема 2.1. *Если* $n = p$, *то в условиях взаимодействия процессов, процессоров и блоков 1–5 минимальное общее время выполнения множества неоднородных конкурирующих процессов* $T_{сн}^{ac}(p, p, s)$ *определяется длиной критического пути в сетевом вершинно–взвешенном графе* G^c, *определяемом матрицей* T^c, *с начальной вершиной* t_{11} *и конечной вершиной* t_{ps}.

Доказательство теоремы следует из построения сетевого вершинно–взвешенного графа G^c, отображающего выполнение p процессов на p процессорах в условиях их взаимодействия 1–5 и свойств критического пути в сетевых графах.

2) $n > p$, т.е. число процессоров многопроцессорной системы меньше числа конкурирующих процессов. Рассмотрим два подслучая: $n = kp$, $k > 1$ и $n = kp + r$, $k \geq 1$, $1 \leq r < p$.

При $n = kp$, $k > 1$, исходную матрицу T^c разбиваем на k подматриц T_m^c, $m = \overline{1, k}$, размерности $p \times s$ каждая. В результате все множество из n процессов будет разбито на k

групп по p процессов в каждой. По каждой из подматриц T_m^c, $m=\overline{1,k}$, строим k линейных диаграмм Ганта, каждая из которых отображает во времени выполнение очередных p процессов на p процессорах.

Очевидно, что если выполнение очередной группы из p процессов начинать только после полного завершения выполнения предыдущей группы, то общее суммарное время выполнения всех n процессов в этом случае будет определяться как сумма длин критических путей в каждой из подряд идущих несовмещенных диаграмм Ганта, задаваемых прямой суммой матриц T_m^c, $m=\overline{1,k}$. Однако это время можно существенно сократить, если воспользоваться приемом совмещения k последовательных диаграмм Ганта по оси времени справа налево. Причем совмещение осуществляется, начиная со второй диаграммы, на максимально возможную величину таким образом, чтобы не нарушались технологические условия выполнения множества процессов и блоков 1–5.

Результирующая матрица T_*^c времен выполнения блоков программного ресурса с учетом максимального совмещения последовательных диаграмм будет состоять из подматриц T_1^c, T_2^c, …, T_k^c типа T^c размерности $p \times s$ каждая. При этом подматрицы T_m^c, $m=\overline{1,k}$, в результирующей матрице T_*^c располагаются таким образом, чтобы не нарушался характер взаимодействия как блоков программного ресурса, выполняемых на одном и том же процессоре (горизонтальные связи между блоками), так и выполняемых на разных процессорах (вертикальные связи) (рис. 2.1).

Первая строка матрицы T_*^c будет состоять из подматриц T_m^c, $m=\overline{1,k}$, что отражает характер взаимодействия блоков программного ресурса и процессов, выполняемых на одних и тех же процессорах.

Каждый шаг совмещения диаграмм определяется соответствующим смещением подматриц T_m^c, $m=\overline{1,k}$, начиная с первой, таким образом, что очередная строка, состоящая из этих подматриц, смещалась справа налево на максимальную величину, не нарушающую линейный порядок выполнения одних и тех же блоков для разных процессов (вертикальные связи). С учетом того, что все подматрицы T_m^c, $m=\overline{1,k}$, имеют одну и ту же размерность $p\times s$, величина смещения на каждом шаге будет равна s. Вместо смещенной на каждом шаге самой правой подматрицы ставятся нули.

Выполнив таким образом $k-1$ шаг совмещений, получим структуру результирующей матрицы T_*^c, соответствующую окончательно совмещенной диаграмме Ганта. Причем в матрице T_*^c будут учтены как все горизонтальные связи между блоками, так и вертикальные, а также связи между блоками из разных диаграмм Ганта. Результирующая матрица T_*^c будет блочной, и, как блочная матрица, она будет симметричной, верхней треугольной относительно второй диагонали, типа Ганкелевой, порядка k, т.е. будет иметь вид:

$$T_*^c = \begin{bmatrix} T_1^c & T_2^c & T_3^c & \cdots & T_{k-1}^c & T_k^c \\ T_2^c & T_3^c & T_4^c & \cdots & T_k^c & 0 \\ T_3^c & T_4^c & T_5^c & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \ddots & \cdots & \cdots \\ T_{k-1}^c & T_k^c & 0 & \cdots & 0 & 0 \\ T_k^c & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Здесь T_m^c, $m=\overline{1,k}$ – $p \times s$ –матрицы вида:

$$T_m^c = \begin{bmatrix} t_{(m-1)p+1,1} & t_{(m-1)p+1,2} & \cdots & t_{(m-1)p+1,s} \\ t_{(m-1)p+2,1} & t_{(m-1)p+2,2} & \cdots & t_{(m-1)p+2,s} \\ \cdots & \cdots & \ddots & \cdots \\ t_{mp,1} & t_{mp,2} & \cdots & t_{mp,s} \end{bmatrix}.$$

Размерность результирующей матрицы T_*^c равна $kp \times ks$.

Далее поступаем аналогично, как и в случае $n \le p$. Строим сетевой вершинно–взвешенный граф с весами, задаваемыми матрицей T_*^c. Вершины этого графа будут расположены в узлах прямоугольной $kp \times ks$ решетки. Очевидно, что веса вершин, соответствующих нулевым значениям матрицы T_*^c, в том числе и конечной вершины с номером (kp, ks), равны нулю.

Тогда, как и в случае $n = p$, имеет место следующая теорема.

Теорема 2.2. *Если $n = kp$, то в условиях взаимодействия процессов, процессоров и блоков 1–5 минимальное общее время выполнения множества неоднородных конкурирующих процессов $T_{сн}^{ac}(p, kp, s)$ определяется длиной критического пути в сетевом* *вершинно–взвешенном графе* G^c, *определяемом матрицей* T_*^c, *с начальной вершиной* t_{11} *и конечной вершиной* $t_{kp,ks}$.

Доказательство теоремы следует из построения сетевого вершинно–взвешенного графа, определяемого матрицей T_*^c, который полностью отображает во времени выполнение $n = kp$, $k > 1$, процессов на p процессорах и, при этом, учитывает все возможные связи между блоками, процессами и процессорами, задаваемыми условиями их взаимодействия в асинхронном режиме, а также свойств критического пути в сетевых графах.

Для второго подслучая, когда число процессов не кратно числу процессоров, т.е. $n = kp + r$, $k \geq 1$, $1 \leq r < p$, матрица времен выполнения блоков структурированного программного ресурса строится аналогично, как и в первом подслучае, с той лишь разницей, что она будет иметь размерность $(k+1)p \times (k+1)s$, а подматрица T_{k+1}^c типа T_m^c будет содержать $p - r$ нулевых строк. Минимальное общее время $T_{сн}^{ас}(p, kp + r, s)$ выполнения множества неоднородных конкурирующих процессов в этом случае также будет определяться длиной критического пути в соответствующем сетевом графе.

Пример 2.1.

Пусть число процессоров многопроцессорной системы $p=3$, число блоков структурированного программного ресурса $s=4$ и число конкурирующих процессов $n=9$. Матрица времен выполнения блоков программного ресурса $T^c=[t_{ij}]$, $i=\overline{1,9}$, $j=\overline{1,4}$, имеет вид:

$$T^c=\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 2 & 1 & 3 & 1 \\ 3 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 1 & 2 & 3 \end{bmatrix}.$$

В нашем примере число конкурирующих процессов больше числа процессоров, более того, число процессов кратно числу процессоров многопроцессорной системы,

$k = 3$. Для этого случая приведены несовмещенная (рис. 2.2) и совмещенная (рис. 2.3) диаграммы Ганта.

Как видно из рис. 2.3, минимальное общее время выполнения всех $n = 9$ неоднородных процессов с учетом максимального их совмещения составляет величину, равную 27 единицам.

Результирующая матрица времен выполнения блоков T_*^c будет состоять из подматриц

$$T_1^c = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 2 & 1 \end{bmatrix}, T_2^c = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 2 & 1 & 3 & 1 \end{bmatrix}, T_3^c = \begin{bmatrix} 3 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 1 & 2 & 3 \end{bmatrix},$$

и иметь следующий вид: $T_*^c = \begin{bmatrix} T_1^c & T_2^c & T_3^c \\ T_2^c & T_3^c & 0 \\ T_3^c & 0 & 0 \end{bmatrix}$, или

$$T_*^c = \begin{bmatrix} \begin{bmatrix} \widehat{1} & \widehat{2} & \widehat{3} & 1 \\ 2 & 3 & \widehat{1} & \widehat{2} \\ 3 & 1 & 2 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 1 & 1 \\ \widehat{2} & 1 & 1 & 2 \\ \widehat{2} & 1 & 3 & 1 \end{bmatrix} & \begin{bmatrix} 3 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 1 & 2 & 3 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 2 & 1 & 3 & 1 \end{bmatrix} & \begin{bmatrix} \widehat{3} & \widehat{2} & 1 & 1 \\ 1 & \widehat{1} & \widehat{3} & 2 \\ 2 & 1 & \widehat{2} & \widehat{3} \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 3 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 1 & 2 & 3 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{bmatrix}.$$

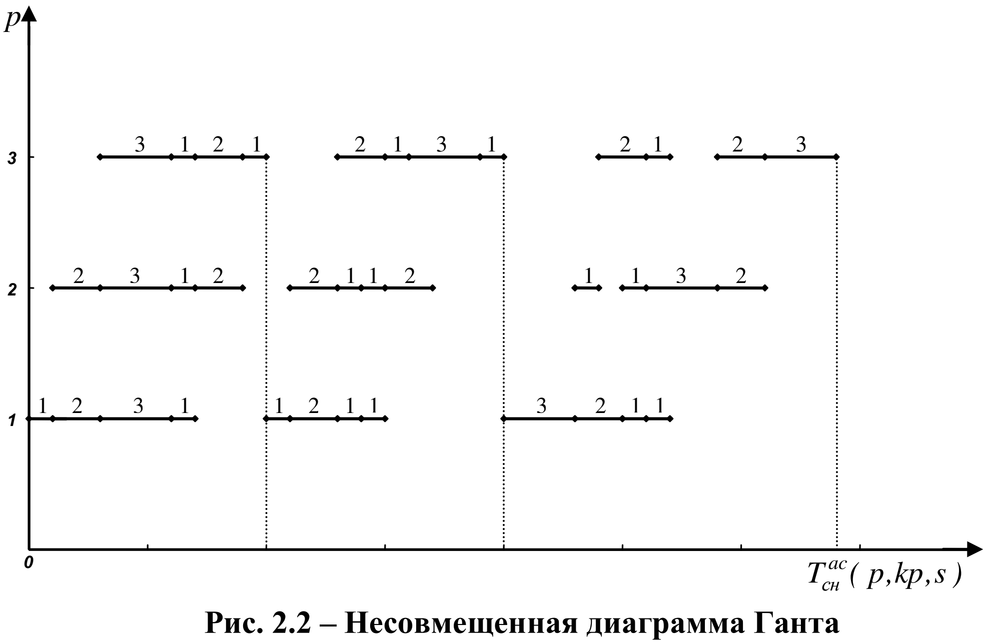

Рис. 2.2 – Несовмещенная диаграмма Ганта

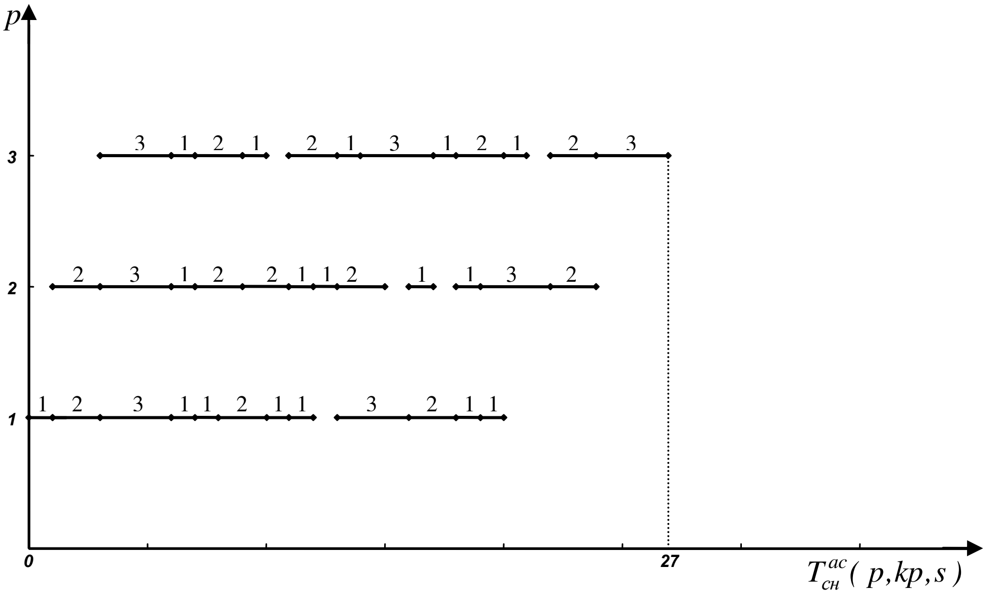

Рис. 2.3 – Совмещенная диаграмма Ганта

Значение критического пути в сетевом графе, определяемом матрицей T^{c}_{*}, равно *27*, что совпадает со значением времени выполнения процессов в совмещенной диаграмме Ганта. Веса вершин, через которые проходит критический путь выделены в матрице символом “$\widehat{x}$”.