ЗАДАЧА ВЫБОРА ЧИСЛА И МЕСТ РАЗМЕЩЕНИЯ ЦЕНТРОВ ХРАНЕНИЯ И ОБРАБОТКИ ИНФОРМАЦИИ В КОМПЬЮТЕРНОЙ СЕТИ

О.В. Есиков,
профессор, доктор технических наук, начальник Управления информационных и аналитических технологий аппарата Администрации Тульской области

Д.В. Изотов,
программист департамента внутренних сервисов ЗАО «Торговый дом «Перекресток», Москва

Адрес: г. Москва, ул. Средняя Калитниковская, д. 28 стр. 4
E-mail: dvikti@km.ru

В статье сформулирована задача выбора числа и мест размещения центров хранения и обработки информации в компьютерной сети по критерию максимума интенсивности поступления запросов на информационное обслуживание. Проведен анализ характера информационных процессов в распределённой компьютерной сети.

Ключевые слова: центры хранения и обработки информации, компьютерная сеть, моделирование информационных процессов, центр графа.

Задача выбора числа и мест размещения центров хранения и обработки информации (ЦХИ) в компьютерной сети (КС) по критерию максимума интенсивности поступления запросов на информационное обслуживание может быть сформулирована следующим образом: требуется определить минимальное число ЦХИ, обслуживающих информационные запросы от автоматизированных рабочих мест должностных лиц (АРМ ДЛ), и такое их размещение в узлах сети, чтобы значение времени задержки передачи сообщения для каждого АРМ ДЛ не превышало допустимой величины, а суммарная приведенная интенсивность поступления запросов на узлы КС, в которых будут размещены ЦХИ, была при этом максимально возможной. Эта задача относится к так называемым задачам о размещении, например, центров скорой помощи, складов и т.п.

Известные математические постановки задач о размещении представляют собой частные случаи классических задач теории графов – «задачи о p-медиане» и «задачи о p-центрах» [1-7], а также «задач о назначении» [8,9], для решения которых, как правило, используются алгоритмы, основанные на идеях метода ветвей и границ. Этим задачам посвящён ряд работ таких авторов, как: S.L. Hakimi, H. Noltermeier, J. Spoerhose, Н. Кристофидес, В.Л. Береснев, Э.Х. Гимади, В.Т. Дементьев, Е.В. Алексеева, Ю.А. Кочетов, Г.Г. Забудский и другие.

В настоящей статье изложен подход к решению задачи, сформулированной применительно к размещению ЦХИ в компьютерной сети на основе известной в теории графов задачи о p-центрах.

С целью формирования математической модели задачи выбора ЦХИ предварительно с применением теории массового обслуживания рассматривается процесс передачи и обслуживания информационных запросов в сети.

Время задержки сообщений в сети можно определить как отрезок времени между моментом начала ввода информации в исходном АРМ ДЛ и моментом получения последнего знака сообщения в узле адресата.

Стохастичность поступления данных и недетерминированный характер их обработки должны быть учтены в ходе моделирования информационных процессов, протекающих в КС. Предположением, необходимым для возможности использования аналитических моделей массового обслуживания в этом случае, можно принять предположение о том, что длительности передачи сообщения (пакета) по разным каналам передачи данных являются независимыми случайными величинами. Для расчета задержек в коммуникационных сетях широкое распространение получила модель сети массового обслуживания, предложенная в [10]. Суть ее состоит в следующем. В коммуникационной сети с коммутацией сообщений (или пакетов) имеется N узлов и W каналов связи, каждый из которых интерпретируются как система массового обслуживания $M/M/1$ (*рис. 1*). И каналы связи, и узлы абсолютно надежные. Пропускная способность w-го канала связи C_w, бит/с, $w=1,\dots,W$.

Узлы выполняют операции по коммутации сообщений, включая их редактирование, выбор маршрутов, буферизацию и т.д. Предполагается, что время обработки в узле постоянно и, более того, пренебрежимо мало. Имеются очереди к каналам связи. При передаче сообщений возникают задержки. В узлы поступает пуассоновский поток запросов, который можно определить как поток между каждой парой узлов КС со средней интенсивностью λ_{ij}, $i, j=1,\dots, N$, $i \neq j$. Объемы сообщений независимы и распределены по показательному закону со средним значением $\frac{1}{\mu}$.

Для размещения этих сообщений в узлах сети имеется память неограниченного объема.

Обозначим λ_w – среднюю интенсивность потока сообщений по w-му каналу связи, тогда

$$\lambda_w = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_{ij} z_{ij}(w),\ w=1,\dots,W,$$

где $\|z_{ij}(w)\|$ – матрица маршрутов передачи данных; $z_{ij}(w)=1$, если при передаче информации из i-го узла сети в j-й она проходит по w-му каналу передачи данных; 0 – в противном случае.

Время задержки сообщений в w-ом канале связи, в соответствии с моделью $M/M/1$, определяется по следующей формуле

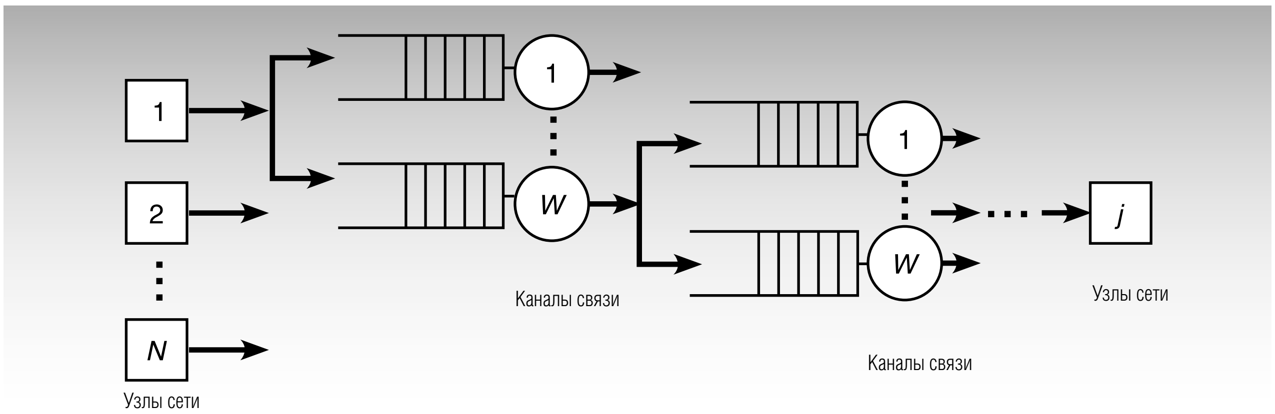

Рис. 1. Сеть массового обслуживания для моделирования информационных процессов в КС.

$$T_w = \frac{1}{\mu C_w - \lambda_w},\ w = 1, ..., W.$$

В реальности технологий обслуживания может быть очень много. Это бесприоритетное и приоритетное обслуживание одним или несколькими приборами, многофазное обслуживание, обслуживание в режиме разделения времени и т.д. Так, например, в [10] показана возможность приближенного расчета задержек в коммуникационных сетях с непоказательными длинами пакетов. Канал связи в этом случае моделируется системой $M/G/1$. Под сообщениями понимаются не только запросы пользователей на получение выходных документов, но и заявки на пересылку файлов или ввод в базу данных информации из входных документов, а также некоторые технологические заявки по управлению вычислительным процессом или контролю над ним.

Для более адекватного отображения функционирования КС и снятия ограничения с учётом введенных выше предположений предлагается использование имитационной модели [11].

Для формализации задачи выбора числа и мест размещения ЦХИ компьютерную сеть представим в виде неориентированного графа $G=(X, Г)$, вершины x_i, $i = 1,..., N$, которого соответствуют узлам сети, а дуги g_w, $w = 1,...,W$ – каналам связи. Тогда «длины» дуг графа G - T_w – образуют матрицу времен задержки передачи сообщения между соответствующими узлами сети (далее для удобства – «время передачи»). Веса q_j, соответствующие вершинам графа, определяют суммарный объем запросов на соответствующие этим вершинам узлы КС.

Для любой вершины x_i графа $G=(X, Г)$ пусть $R^o(x_i)$ есть множество тех вершин x_j графа G, которые достижимы из вершины x_i с помощью путей с взвешенными «длинами» $q_jT(x_i, x_j)$, не превосходящими величины T_{max}, где $T(x_i, x_j)$ – «длина» кратчайшего пути от вершины x_i до вершины x_j.

Через $R^t(x_i)$ обозначим множество тех вершин графа G, из которых вершина x_i может быть достигнута с использованием путей, имеющих взвешенные длины $q_jT(x_j, x_i) \geq T_{max}$.

Таким образом,

$$R^o(x_i) = \{x_j \mid q_jT(x_i, x_j) \leq T_{max},\ x_j \in X\} \text{ и}$$
$$R^t(x_i) = \{x_j \mid q_jT(x_j, x_i) \leq T_{max},\ x_j \in X\}. \quad (1)$$

Для каждой вершины x_i, соответственно из множеств $R^o(x_i)$ и $R^t(x_i)$, определим следующие два числа:

$$s_o(x_i) = max[q_jT(x_i, x_j)] \text{ и}$$
$$s_t(x_i) = max[q_jT(x_j, x_i)]. \quad (2)$$

Числа $s_o(x_i)$ и $s_t(x_i)$ называются соответственно числом внешнего разделения и числом внутреннего разделения вершины x_i.

Если T_0 – наименьшая «длина» T, такая, что для вершины x_i

$$R^o(x_i) = X,$$

(т.е. все вершины графа G достижимы из x_i с использованием путей, взвешенные «длины» которых не превосходят T_0, причем T_0 – наименьшее из таких чисел), то из соотношений (1) и (2) следует равенство

$$s_o(x_i) = T_0.$$

Аналогично, если T_t – такая наименьшая длина T, что

$$R^t(x_i) = X, \text{ то}$$
$$s_t(x_i) = T_t.$$

Вершина x_o, для которой

$$s_o(x_o^*) = min[s_o(x_i)],$$

(т.е. вершина, соответствующая минимальному из всех чисел $s_o(x_i)$, ранее определённых по указанному выше правилу) называется внешним центром графа G.

Аналогично вершина x_t, для которой

$$s_t(x_t^*) = min[s_t(x_i)]$$

называется внутренним центром графа G.

Понятие центра графа допускает следующее обобщение: можно рассматривать не отдельную точку (центр), а множество из p точек, которые образуют кратный центр (p-центр).

Пусть X_p – подмножество (содержащее p вершин) множества X вершин графа $G = (X, Г)$. Через $T(X_p, x_i)$ будем обозначать наикратчайшее из расстояний между вершинами множества X_p и вершиной x_i, т.е.

$$T(X_p, x_i) = min\ [T(x_j, x_i)],\ x_i \in X_p,\ x_j \in X;$$

Аналогично

$$T(x_i, X_p) = min\ [T(x_i, x_j)],\ x_i \in X_p,\ x_j \in X.$$

Подобно тому, как определялись числа разделения вершин, можно определить числа разделения для множества вершин:

$$s_o(X_p)= max\,[q_j T(X_p, x_j)] \text{ и}$$
$$s_t(X_p)= max\,[q_j T(x_j, X_p)], \text{ где}$$

$s_o(X_p)$ и $s_t(X_p)$ – числа внешнего и внутреннего разделения множества X_p.

Множество $X_{po}{}^*$, для которого

$$s_o(X_{po}{}^*)= min\,[s_o(X_p)],$$

называется p-кратным внешним центром графа G; аналогично определяется p-кратный внутренний центр $X_{pt}{}^*$ [1].

Следует отметить, что числа внутреннего и внешнего разделения множества X_p рассматриваются при условии достижимости любой вершины графа из множества X_p, что вполне возможно в случае решения задачи для неориентированного графа, соответствующего распределенной компьютерной сети с допустимой передачей данных по каналам связи во всех направлениях.

Очевидно, что внешний и внутренний центры неориентированного графа совпадают, так как в этом случае числа разделения $s_o(X_p)$ и $s_t(X_p)$ равны между собой для любого множества X_p.

Таким образом, исходя из вышеизложенного, задача выбора числа ЦХИ и их размещения в узлах сети будет состоять в нахождении p-центров соответствующего графа G для различных значений p до тех пор, пока число разделения p-центра не станет меньше или равно заданной величине. Полученное (последнее) значение числа p будет наименьшим числом ЦХИ, а p-центр – их оптимальным размещением, удовлетворяющим предъявляемым требованиям.

Исходя из выше изложенного, общая задача определения p-центра применительно к выбору ЦХИ в распределенной компьютерной сети может быть сформулирована следующим образом [12].

Для заданного «критического» расстояния найти такое наименьшее число центров и такое их размещение, чтобы все вершины графа лежали в пределах этого критического расстояния (по крайней мере, каждая вершина – от ближайшего к ней центра).

Очевидно, что центры графа легко могут быть получены из матрицы весов дуг графа. Однако находить полным перебором p-центр можно лишь для небольших графов и для небольших значений величины p. При таком подходе надо построить всевозможные множества вершин $X_p \subseteq X$, содержащие p вершин, а затем непосредственно найти множества $X_{po}{}^*$ и $X_{pt}{}^*$. Этот процесс потребует выполнить

$$p \times (N\text{-}p) \times \binom{N}{p}$$

сравнений, что даже при небольших значениях N и p неприемлемо много.

Для нахождения p-центра существует ряд эвристических методов, например, метод Сингера [1]. Они имеют низкую эффективность по быстродействию.

Для уменьшения числа переборов предлагается метод, основанный на сведении задачи о p-центрах к задаче о покрытии. Идея такого метода состоит в следующем.

На предварительном этапе матрица весов дуг преобразуется в матрицу $\|d_{ij}\|$, элементами которой являются кратчайшие пути между всеми парами вершин графа G. Затем, исходя из полученной матрицы $\|d_{ij}\|$, составляется бинарная матрица (матрица покрытий) $\|a_{ij}\|$, каждый элемент которой равен 1, если i-я вершина достижима из j-й вершины в пределах критического расстояния, т. е. если $d_{ji} \le T_{max}$, и 0 – в противном случае. На заключительном этапе решается задача о покрытии, исходными данными для которой будут являться матрица покрытий $\|a_{ij}\|$ и веса q_i, приписанные вершинам графа G, а результатом – искомый p-центр.

Рассмотрим математическую формулировку задачи о покрытии, применительно к которой можно выбрать достаточной эффективный метод решения.

Используя матрицу кратчайших путей $D = \|d_{ij}\|$, перейдем к матрице покрытий $A=\|a_{ij}\|$, которая составляется по следующим правилам:

$$a_{ij} = \begin{cases} 1,\ если\ d_{ij} \le T_{max} \\ 0,\ в\ противном\ случае\ . \end{cases}$$

Задача о нахождении p-центра на графе, удовлетворяющего заданным условиям, сводится к нахождению такого наименьшего множества X^*, чтобы из каждой вершины графа была достижима в пределах заданного расстояния T_{max} хотя бы одна вершина, входящая в множесво X^*, т.е. чтобы

$$\sum_{i \in X^*} a_{ij} \ge 1\ ,\ \ j=1,\dots,N\ .$$

Сумма $\sum_{i \in X^*} q_i$ должна быть при этом максимально возможной (для обеспечения минимума передаваемой в сети информации).

Таблица 1.

№ уз-ла	1	2	3	4	5	6	7	8	9	10	11	12	13	14	q
1	0	25	6	12	11	4	14	17	21	22	23	24	20	19	36
2	25	0	7	10	12	6	10	10	11	14	15	16	17	11	18
3	6	7	0	7	5	5	9	5	5	7	8	9	10	12	5
4	12	10	7	0	4	3	2	4	3	4	8	7	6	5	6
5	11	12	5	4	0	5	6	7	7	8	9	9	8	7	6
6	4	6	5	3	5	0	3	4	7	8	9	12	18	13	6
7	14	10	9	2	6	3	0	8	8	9	10	10	9	8	9
8	17	10	5	4	7	4	8	0	4	5	7	9	9	8	10
9	21	11	5	3	7	7	8	4	0	7	8	9	7	6	12
10	22	14	7	4	8	8	9	5	7	0	6	11	12	9	12
11	23	15	8	8	9	9	10	7	8	6	0	17	12	13	12
12	24	16	9	7	9	12	10	9	9	11	17	0	11	10	12
13	20	17	10	6	8	18	9	9	7	12	12	11	0	12	6
14	19	11	12	5	7	13	8	8	6	9	13	10	12	0	9

Эта задача может быть сформулирована в виде задачи целочисленного линейного программирования с булевыми переменными следующим образом.

Требуется максимизировать целевую функцию

$$Q=\sum_{i=1}^{N} q_i x_i \qquad (3)$$

при ограничениях

$$\sum_{i=1}^{N} a_{ij} x_i \geq 1, \ j=1,\ldots, N , \qquad (4)$$

$$x_i=\{0;1\}, \ i=1,\ldots,N , \qquad (5)$$

$$q_i \geq 0 ;$$

$$a_{ij}=\begin{cases} 1, \text{ если } d_{ij} \leq T_{max} \\ 0, \text{ в противном случае} . \end{cases}$$

Задача (3) – (5) относится к широкому классу задач, которые носят название задач о покрытии. Она является задачей целочисленного программирования с булевыми переменными и, следовательно, может быть решена общими методами целочисленного программирования [13].

Из теории алгоритмов известно, что задача о покрытии относится к классу универсальных переборных задач, или NP-полных задач [14]. Для её решения до сих пор неизвестны, а возможно – и не существуют, полиномиальные алгоритмы. Для всех известных алгоритмов решения универсальных переборных задач время счета растет экспоненциально с ростом размерности задачи. Однако характер экспоненциальной зависимости существенно зависит от особенностей каждого алгоритма.

В [12] предлагается метод ветвей и границ. Специфические особенности задачи (3) – (5) позволяют значительно упростить процедуру вычисления нижней границы решения на основе использования двойственной задачи [15].

Рассмотрим решение задачи по критерию максимума интенсивности поступления запросов на узлы сети. С помощью имитационного моделирования получены следующие исходные данные, которые приведены в *табл. 1*, где содержатся времена задержек (в секундах) информации при ее передаче из одного узла сети в другой. Значения q представляют собой «веса» узлов сети.

В *табл. 2* приведены результаты решения данной задачи, где для различных значений заданного времени задержки сообщения в сети указаны номера узлов сети, в которых размещаются ЦХИ (p-центр графа).

Таблица 2.

Заданное время задержки передачи сообщения в сети, с	Номера узлов сети, в которых размещаются ЦХИ (р-центр)
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
2	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
3	1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14
4	1, 2, 3, 4, 11, 12, 13, 14
5	1, 2, 4, 8, 11, 12
6	1, 2, 7, 8
7	1, 2, 8
8	2, 8
9	2, 8
10	8
15	1
20	1

Полученные данные позволяют определить оптимальную, с точки зрения выбранного критерия, конфигурацию сети для заданного времени задержки передачи сообщения, либо для заданного количества ЦХИ и их размещения в сети определить, какая будет при этом максимальная задержка сообщения. ■

Литература

1. Кристофидес Н. Теория графов. Алгоритмический подход. – М: Мир, 1978. – 432 с.
2. Алексеева Е.В., Орлов А.В. Генетический алгоритм для конкурентной задачи о р-медиане // Труды 14 Байкальской международной школы-семинара «Методы оптимизации и их приложения». Том 1. – Северобайкальск: 2008. - с. 570-585.
3. Забудский Г.Г., Филимонов Д.В. Решение дискретной минимаксной задачи размещения на сети // Известите вузов. Математика, № 5 – Омск: 2004. - с. 33-36.
4. Кочетов Ю.А., Кононов А.В., Плясунов А.В. Конкурентные модели размещения производства // Журнал вычислительной математики и математической физики. Т. 49, № 6. - 2009. - с. 1-17.
5. Hakimi S.L. On locating new facilities in a competitive environment. European J. Oper. Res. 1983. V. 12, P. 29–35.
6. Noltermeier H., Spoerhose J., Wirth H.C. Muliple voting location and single voting location on trees. European J. Oper. Res. 2007. V. 181. P. 65-667.
7. Spoerhose J., Wirth H.C. (r,p)-Centroid problems on paths and trees. Tech. Report 441, Inst. Comp. Science, University of Würzburg, 2008.
8. Береснев В.Л. , Гимади Э.Х., Дементьев В.Т. Экстремальные задачи стандартизации. - Новосибирск: Наука, 1978. - 335с.
9. Гимади Э.Х. Обоснование априорных оценок качества приближенного решения задачи стандартизации // Управляемые системы. - Новосибирск, 1987. - Вып. 27. – с. 12-27.
10. Клейнрок Л. Вычислительные системы с очередями: Пер. с англ. - М.: Мир, 1979.- 600 с.
11. Акиншин Н.С., Подчуфаров Ю.Б., Изотов В.Н., Комогорцев П.В. Имитационная модель для решения задачи синтеза физической структуры АСУ нового поколения. //Оборонная техника. 1997. № 3-4. - С. 99-103.
12. Изотов Д.В. Оптимальное размещение центров хранения и обработки информации по критерию максимума интенсивности запросов / Е.В. Ларкин, Д.В. Изотов // Журнал «Прикладная информатика», № 3 (33). – М: Изд-во ООО «Маркет ДС Корпорейшн». – 2011. – с. 37 - 42.
13. Финкелыптейн Ю.Ю., Корбут А.А. Дискретное программирование. - М.: Наука, 1969. - 368 с.
14. Современное состояние теории исследования операций / Под ред. Н. Н. Моисеева. - М: Наука, 1979. – 303 с.
15. Алексеев О. Г., Григорьев В. Ф. Некоторые алгоритмы решения задачи о покрытии и их экспериментальная проверка на ЭВМ // ЖВМ и МФ. - 1984. - Т.24, №10. - с. 1565-1570.