Занятие №3

Исследование характеристик ИС, обрабатывающих заявки по заданной технологии

Теоретические сведения

По архитектуре информационная система (ИС) имеет вид сложного производственного комплекса, включающего в себя вычислительные средства, математическое и информационное обеспечение, периферийное оборудование, людей. Сложность ИС порождает проблему оптимальной организации системы. Для ее успешного решения необходимо уметь строить модели ИС. Практический интерес моделирование имеет в двух ситуациях:

1) при анализе существующей системы с целью ее модернизации и предварительной оценки результатов модернизации;
2) при анализе вариантов построения новой системы с целью выбора лучшего варианта.

Критерием для разрешения проблемной ситуации как при модернизации, так и при проектировании системы служит ее производительность и (или) среднее время реализации процесса обработки одного задания — цикл обработки заявок. Поэтому результатом моделирования должно быть среднее время реализации процесса и пропускная способность системы как функция параметров системы и процесса.

Особенности ИС.

Сложность системы (десятки или сотни элементов: устройств и людей), вероятностный характер входящего потока заявок на проектирование и времени выполнения отдельных операций проектирования свидетельствуют о том, что задача определения среднего времени реализации процесса обработки заданий есть задача теории сетей СМО или теории стохастических сетей.

Время пребывания требования в любой сети

$$T = T_n = \sum_i \alpha_i T_i \ ,$$

где T_i - среднее время пребывания требования в элементе i; α_i -среднее число передач, значения α_i находятся решением системы линейных уравнений:

$$\lambda_j = \sum_i \theta_{ij} \lambda i, ij = \overline{0,n}, \alpha_i = \frac{\lambda_i}{\lambda_0},$$

где λ_0 - интенсивность входящего потока в сеть; λi - интенсивность входного потока на элемент i; θ_{ij} - вероятность передачи заявки из элемента i после ее обслуживания.

Для определения цикла обработки необходимо знать величины T_i и θ_{ij}. Кроме того, они являются исходными данными для подбора различных элементов системы. Каждое требование инициирует решение определенной последовательности задач проектирования. В обычной стохастической сети все требования независимы. В ИС требования также независимы, поскольку внутри сети требование выступает как фаза в реализации процесса, как требование на решение одной задачи из последовательности, т.е. требования выступают как фазы разных процессов и, следовательно, независимы. Однако взаимосвязь задач каждого отдельного процесса налагает условия на переход требования из одного узла в другой. В стохастической сети это отображается весами дуг связей узлов сети. Оценки весов дуг вероятностей передач θ_{ij} - от узла к узлу - должны быть получены как исходные данные для анализа системы. Так как последовательность задач не детерминирована и по разным причинам отдельные подмножества задач приходится решать повторно, иногда не один раз, то это ведет к тому, что для разных требований фактическая последовательность задач неодинакова. Кроме того, разные задачи могут решаться одним элементом системы и, наоборот, одна задача - сразу несколькими элементами. Поэтому прямая оценка вероятностей передач θ_{ij} практически затруднена.

Следует отметить и такую особенность ИС по сравнению с ВС. При моделировании ВС люди, т.е. коллектив пользователей, выступают как источники требований. Их поведение отображается в моделях потоков требований. В ИС люди решают часть задач из общей последовательности, следовательно, принципиально не устранимы из системы и должны быть представлены в модели системы как ее элементы, т.е. как СМО.

Процесс может быть представлен с разной степенью детализации.

При моделировании важно знать, какими устройствами или людьми выполняются отдельные операции, поэтому разбиение процесса на операции следует производить, руководствуясь логикой работы систем и людей. Разумно выделять в качестве операций такие части процесса, которые реализуются без прерывания, без обращения к другим элементам одним устройством или одним человеком. В то же время при диалоговом решении какой-либо задачи не всегда необходимо отделять операции человека и ЭВМ. При моделировании, процесса в целом за операцию может быть принят весь цикл взаимодействия человека и ЭВМ при решении задачи.

Расчленение процесса на операции всегда неформально. В зависимости от цели и искусства человека могут быть построены разные схемы для изображения одного и того же процесса.

Последовательность операций может быть представлена в виде графа G_p, узлы которого отображают операции, а дуги - связи между операциями. В каждый данный момент времени набор операций фиксирован и известно, какими элементами системы (устройствами или людьми) каждая операция реализуется. Оценка среднего времени выполнения данной операции данным элементом так же, как и оценка частоты перехода от данной операции к другим, не вызывают тех трудностей, которые возникают при получении оценок параметров системы. Если характеристики процесса получены, то процесс можно представить в виде дважды взвешенного графа: веса узлов - оценки времени выполнения операций t_k веса дуг - оценки вероятностей P_{kl} переходов от операции к операции.

Связь с параметрами системы осуществляется косвенно: оценки времени выполнения операций t_k сделаны относительно конкретных элементов системы, т.е. фактически дано распределение операций по элементам. Это можно зафиксировать в виде матрицы D распределения операций по элементам. Элемент $d_{ki} = 1$, если *k*-я операция реализуется *i*-м элементом, в противном случае $d_{ki} = 0$. Вероятности P_{kl} могут быть записаны в виде матрицы $[P_{kl}]$, $k,l = \overline{1,k}$, а оценки среднего времени выполнения операций - в виде .диагональной матрицы $[t_k]$.

В целом можно сказать, что моделью процесса является сеть СМО, в которой каждая отдельная СМО служит моделью отдельной операции. При соответствующих предположениях о законах распределения времени выполнения операций могут быть вычислены характеристики сети. В частности, может быть вычислено среднее время реализации процесса как среднее время пребывания требования в сети. В качестве требований внутри сети выступают требования на выполнение отдельных операций.

Вычисление параметров модели системы производится по модели процесса

Получение параметров системы T_i, θ_{ij}, α_i по характеристикам процесса осуществляется следующим образом.

1. Прежде всего определяем, сколько раз реализуется операция к за одну реализацию процесса. За счет возвращений от последующих операций к предыдущим процесс представляется реализацией некоторого пути в вероятностной структуре. Число реализаций операций *k* может быть лишь средней оценкой по многим процессам. Получить ее можно, опираясь на характер матрицы $[P_{kl}]$. Вслед за выполнением операции *k* обязательно должна выполняться другая операция, иначе заявки уходят из сети. Ввиду этого сумма $\sum_{l=1}^{k} P_{kl} = 1$. А это значит, что матрица описывает некоторую марковскую цепь. И сама задача определения среднего числа реализаций операции *k* сводится к известной задаче теории стохастических сетей. Считая, что λ_0 -интенсивность входящего потока, и учитывая, что

$$\lambda_i = \sum_{k=0} P_{kl}\lambda_{k,} \quad l = \overline{0,k},$$

находим $\frac{\lambda_k}{\lambda_0} = \alpha_k$ - среднее число реализаций операции *k* за одну реализацию процесса. Реальные режимы системы таковы, что входящий поток заявок определяется не внешними источниками, а производительностью самой системы. Поэтому принимаем замкнутую модель. Для этого выход системы надо замкнуть связью с весом (равным I) на нулевой элемент, который таким образом превращается в фиктивный элемент системы с нулевым временем обслуживания. Если λ_0 не задана, то определитель системы равен 0, поэтому она не имеет однозначного решения. Однако можно однозначно определить отношения $\frac{\lambda_k}{\lambda_0} = \alpha_k$.

Из полученных таким образом величин α_k формируется диагональная матрица $[\alpha_k]$.

2. Определяем среднее число операций α_i, реализуемых на элементе i. Величина α_i аналогична α_k, но α_k - это параметр процесса, а α_i - параметр системы. Содержательно процесс получения α_i можно пояснить следующим образом. Предположим, что на элементе i реализуются операции *k,l,m*. Эти операции реализуются соответственно $\alpha_k, \alpha_l, \alpha_m$ раз за одну реализацию процесса. Общее число операций, реализуемых на элементе i за одну реализацию процесса, равно сумме $\alpha_k + \alpha_l + \alpha_m$. Распределение операций по элементам представлено матрицей D. Поэтому

$$[\alpha_i] = D'[\alpha_k]D,$$

где $[\alpha_i]$ - диагональная матрица из α_i.

3. Определение среднего времени T_i реализации операции элементом i. Пусть на элементе с реализуются операции *k,l,m*. Известно, какое число раз каждая из них реализуется $\alpha_k, \alpha_l, \alpha_m$, кроме того, известны и оценки времени их выполнения t_k, t_l, t_m. T_i может быть лишь средним временем, сформированным из указанных величин, т.е.

$$T_i = \frac{t_k a_k + t_l a_l + t_m a_m}{a_k + a_l + a_m}$$

Сумма в знаменателе равна α_i. Учитывая диагональный характер матрицы $[\alpha_i]$ знаменатель можно получать как элемент матрицы, обратной $[\alpha_i]$, т.е. $\left[D^T[\alpha_k]D\right]^{-1}$. В числителе стоит сумма произведений $\alpha_k t_k$, причем в сумму входят лишь те члены, которые соответствуют операциям, реализуемым на элементе *i*. Поэтому числитель можно получить как элемент матричного произведения $\left[D^T[\alpha_k][t_k]D\ \right]$ в целом получаем

$$[T_i] = \left[D^T[\alpha_k]D\right] \bullet \left[D^T[\alpha_k][t_k]D\right],$$

где $[T_i]$- диагональная матрица.

4, Определение вероятностей передач θ_{ij} от элемента i к элементу j. Передача от i-го элемента элементу j происходит всякий раз, когда за операцией, реализуемой на элементе *i*, следует операция, реализуемая на элементе *j*. Поскольку связь операций не детерминированная, а вероятностная, выражением чего служит матрица $[P_{kl}]$, то и передачи от элемента к элементу происходят с некоторой вероятностью. Пусть на элементе *i* реализуются операции *(k,l,m)*, причем за операциями *k* и *l*следуют с вероятностями $P_{kk'}$ и $P_{ll'}$ операции *k'* а *l'*, которые реализуются на элементе *i*, а за операцией *m* подобной операции не следует. На элементе *i* реализуется $(\alpha_k + \alpha_l + \alpha_m)$ операций за одну реализацию процесса, из них $(a_k P_{kk'} + a_l P_{ll'})$ число раз следующая операция будет выполняться элементом *j*, т.е. вероятность передачи от элемента *i* к элементу *j* равна:

$$\frac{\alpha_k P_{kk'} + \alpha_k P_{ll'}}{(\alpha_k + \alpha_l + \alpha_m)}$$

А вероятность передачи от элемента *i* к элементу *j* после выполнения операции k равна $\dfrac{\alpha_k P_{kk'}}{(\alpha_k + \alpha_l + \alpha_m)}$.

Аналогичным образом находятся вероятности передач от элемента *i* к другим элементам после выполнения операции *k*, а также и операций *l* и *m*. Поскольку $\sum_{n=1}^{k} P_{kn} = 1$ и аналогично $\sum_{n=1}^{k} P_{\ln} = 1$, $\sum_{n=1}^{k} P_{mn} = 1$, и каждое из P_{kn} взвешивается величиной α_k, а $P_{\ln}$ и P_{mn}- величинами α_l и α_m соответственно, то $\sum_{j=1}^{N} \theta_{ij} = 1$. Для получения знаменателя $(\alpha_k + \alpha_l + \alpha_m)$ можно использовать опять же обратную матрицу $[D^T[\alpha_k]D]^{-1}$, а числители получить как элементы следующего матричного произведения: $[D^T[\alpha_k][t_k]D\ ]$. В итоге формула для вычисления θ_{ij} имеет вид

$$[\theta_{ij}] = [D^T[\alpha_k]D]^{-1}[D^T[\alpha_k][t_k]D\ ],$$

где $\lfloor \theta_{ij} \rfloor$ - матрица вероятностей передач системы.

Рассмотрим конкретный пример…

Информационные системы обработки заявок по заданной технологии.

Существует технологический процесс обработки информации – например АСУТП (автоматизированная система управления технологическим процессом).

Примеры таких систем – производственная система, системы автоматизированного проектирования (САПР), экономические системы.

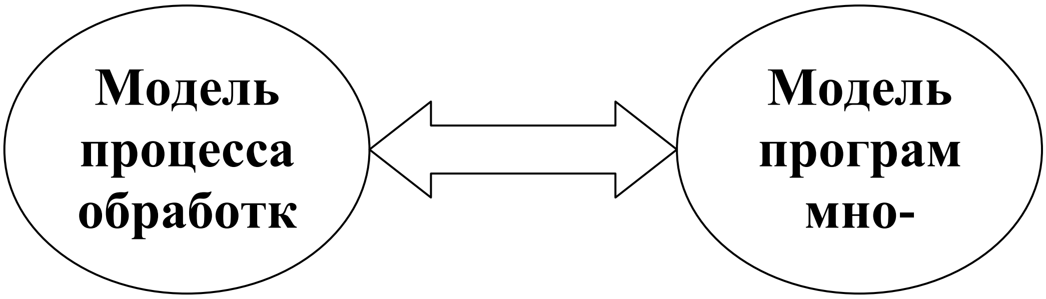

Рис. 3.1 Двойственные информационные системы

Модель процесса обработки информации.

Модель процесса обработки информации представляет собой дважды взвешенный граф (рис. 3.2)

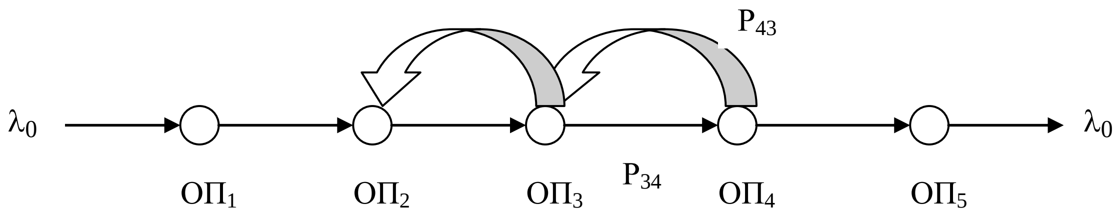

Рис. 3.2 Процесс обработки информации

Ребра показывают пути обработки информации.

$t_{ОПi}$ – время выполнения *k*-ой операции на *i*-ом компоненте.

Каждая операция может выполняться на нескольких типах компонентов, но в процессе обслуживания используется только один компонент.

На каждой операции *k* существует перечень компонентов, на которых она может выполняться, и время ее выполнения:

$$ОП_k \Longleftrightarrow \begin{matrix} k_1 \\ k_2 \\ \\ k_i \end{matrix} \Longleftrightarrow \begin{matrix} tоп_1 \\ tоп_2 \\ \\ tоп_i \end{matrix}$$

Весом вершины является время выполнения операции.

Весами ребер являются вероятности передачи заявок от одной операции к другой P_{kl}.

$\sum P_{kl} = 1$

Приведенный процесс реализуется на системе из трех компонентов (рис. 3.3):

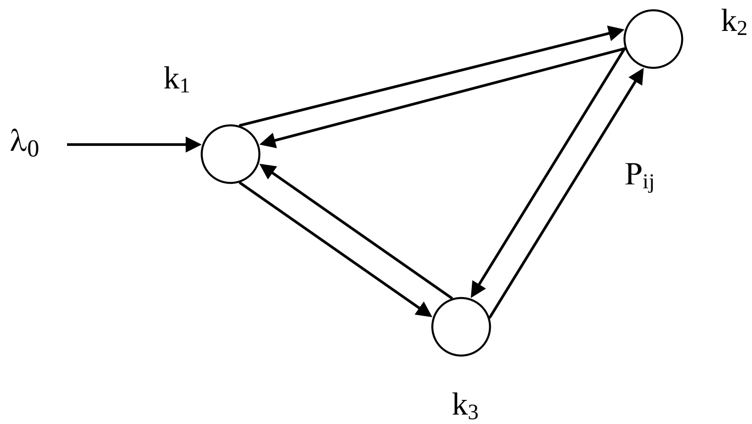

Рис. 3.3 Выполнение операций на системе из 3-х компонентов

(i, j) – номера компонентов технических средств.

компонент i

D_{ki} =

1			
2		d_{ki}	
3			
4			
5			

операция *k*

Строки соответствуют операциям, столбцы соответствуют компонентам системы.

Величина $d_{ki} = 1$, если *k*-я операция реализуется на *i*-м компоненте.
В каждой строке только одна единица, но в столбце их может быть несколько, то есть на одном компоненте выполняется несколько операций.