

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ И СКОРОСТИ ОБРАБОТКИ ДАННЫХ

Умаров А.А.¹, Тукубаев З.Б.².

(Международный Казахско-Турецкий университет имени
Ясауи, г.Туркестан, Казахстан)

В статье рассматривается задачи оптимизации распределения потоков по группам пользователей и скорости обработки информации. Приведен алгоритм решения задачи с помощью ПО Excel с использованием метода динамического программирования.

Ключевые слова: организационные системы, корпоративные информационные системы, консалтинговые компании, анализ бизнес-процессов компании, информационные системы корпоративного управления.

1. Введение

Одной из приоритетных направлений, является **автоматизация** деятельности государственного аппарата, учреждений образования и культуры, ускорения реализации проекта "электронного правительства" [1,2]. Стоит подчеркнуть, что в разделе *Цели – Показатели* Стратегического плана развития вуза МКТУ им. А. Ясауи на 2009-2013 гг. [3, с. 72] определена конкретная цель – **разработка информационной системы управления вузом**.

Задача разработки и внедрения информационных систем (ИС) корпоративного управления состоит из нескольких этапов [4,5,6]. Как правило, речь идет о внедрении в вуз масштабных систем типа ERP (MRP) или прикладных систем, рассчитанных на большую организационную систему. Каждая компания – разработчик корпоративных систем имеет свой внутренний

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru) 1
2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

стандарт внедрения таких систем для каждой конкретной компании. Значительное число консалтинговых компаний специализирующихся на анализе, подборе и внедрении корпоративных информационных систем имеют технологию настройки и запуска системы.

Среди основных задач при разработке ИС можно выделить [6]:

- 1) изучение и анализ бизнес-процессов компании, фиксация бизнес-процессов, реинжиниринг процессов, если необходимо [14];
- 2) определение ключевых пользователей (автоматизированных рабочих мест) [14], их функции, обучение персонала в системе конкретных функций;
- 3) определение оптимального распределения потоков и оптимальной скорости обработки для настройки конфигурации системы;
- 4) внесение необходимых изменений в конфигурацию и настройки системы на основании результатов опытной эксплуатации; определение окончательного списка пользователей системы и фиксация ролей; запуск системы в промышленную эксплуатацию.

2. Основная часть

Задача исследования потоков заявок, поступающих на веб - сервер организации, в частности на центральный сервер МКТУ им. Ясауи описана в [16,17,18]. В работе был определен пуассоновский закон поступления заявок. Результаты, полученные в последних работах дают основу для настоящей задачи.

Настоящая работа посвящается решению задачи 3 и состоит из двух подзадач:

- оптимизация распределения потоков по группам пользователей;
- оптимизация скорости обработки данных.

Эта задача на языке системного аналитика выглядит так; на основании качественного анализа характера транзакции /связей между ними, их типов и количеств/, которые обрабатываются ИС, определяются типы и параметры аппаратного обеспечения; будут определены используемые модули и основные параметры программного обеспечения, а также необходимое количество задействуемых специалистов.

Существуют стандартные методы, позволяющие оценить параметры аппаратного обеспечения для достижения требуемого уровня быстродействия и надежности обработки данных. Привлечение экспертов по ИС позволяет определить, какую систему и какие конкретно модули необходимо использовать, каковы должны быть параметры базовой конфигурации. Необходимо отметить, что независимо от уровня компетенции специалистов однозначного ответа на этот вопрос не существует. При этом, возможно несколько альтернативных решений. Один из возможных подходов к выбору базовой конфигурации описан в [6]. Модели исследования производительности технических систем – сети, серверов рассмотрены в [9, 11, 12, 13, 18].

В конечном итоге, когда бизнес-процессы описаны, а также определены конкретная система и ее модели, возникает задача: как распределить элементарные транзакции между K исполнителями (ключевыми пользователями и/или отдельными подразделениями). Критерием оптимальности такого распределения обычно является *поток информации* (нагрузка и сложность операции), приходящие на одного исполнителя (задача 1), а также минимизация общего *времени обработки информации* (задача 2). В дальнейшем будем использовать максимально допустимый объем информации на одного исполнителя в единицу времени (в сек.)

Постановка задачи 1.

После анализа бизнес-процессов вуза можно определить понятие сложности бизнеса. Пусть всего в бизнес-процессах встречается N операций (элементарных транзакции). На

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru) 3
2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

основании накопленной статистики за квартал или год можно каждой операции t_j сопоставить относительную частоту ее появления v_j , которая при выборке и устойчивых условиях деятельности будет стремиться к вероятности p_j . Пример такого распределения группам пользователей (отделам) представлен на рис. 1 для практического случая внедрения в вузе.

На основании формулы для количества информации $I = k \ln W$ можно определить сложность и количество информации, приходящее на одну элементарную презентацию. Коэффициент k в дальнейшем будем полагать равным 1, что соответствует выбору в качестве шкалы измерения бинарной системы. Как правило, в стандартных ИС каждой транзакции соответствует одна или несколько записей, где каждая запись включает l полей. Совокупность значений всех полей полностью определяет содержание транзакции. Предположим, что значения полей независимы и принимают $\Omega = \{w_1, w_2, \dots, w_l\}$ значений. Тогда полное количество независимых $W_j = w_1 \cdot w_2 \cdot \dots \cdot w_l$ состояний для транзакции данного типа равно. Соответственно количество информации $I_j = k \ln W_j$ - определяет сложность транзакции данного типа t_j .

Таблица 1. Статистика транзакции

Транзакции	Ректорат	Учотдел	Метод. Отдел	Научотдел	Библиотек	Деканат	Преподаватели (10 фак)	Студенты (10 тыс)	Все го	В %
ПОИСК	51	1334	409	420	21	2861	4876	4025	5088	8,75

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru)
2. Зухирхан Бейсекович Тукубаев, кандидат технических наук (zuhr@pochta.ru)

студента по ФИО	3				9			6	8	
ПОИСК студента по спец.	461	1211	309	313	0	2603	4690	32302	41889	7,20
ПОИСК студента по группе	383	318	285	276	0	2777	3902	38184	46125	7,93
ПОИСК преподавателя по ФИО	525	546	2308	2211	0	2344	7954	38415	54303	9,34
ПОИСК препод. по предмету	320	493	1869	1624	0	1990	7112	36068	49476	8,51
ПОИСК препод. по фак/каф	363	417	1960	1883	0	2203	6612	34413	47851	8,23
ПОИСК номера приказа	1068	436	389	410	376	566	1003	0	4248	0,73
ПОИСК номера исх/вх. Докумен.	1103	580	406	415	390	572	1022	0	4488	0,77
РАСЧЕТ колич. исх/вх. Докумен.	501	380	311	295	251	244	0	0	1982	0,34
РАСЧЕТ колич. обработ. Докумен. /завершенных дел	486	566	330	399	283	233	0	0	2297	0,39
РАСЧЕТ колич. студ. по вузу в разрезе на факульт.	0	365	0	0	0	1862	2066	9311	13604	2,33
РАСЧЕТ баллов GPA студ. по	0	482	0	0	0	1607	3991	13022	19102	3,28

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru)

2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

группе										
РАСЧЕТ общей нагрузки часов по факульт.	0	885	0	0	0	2334	2668	0	5887	1,01
РАСЧЕТ нагрузки препод вуза	0	742	0	0	0	3457	3435	0	7634	1,31
РАСЧЕТ общего расписан ия занятий по факульт.	0	611	0	0	0	2360	0	0	2971	0,51
РАСЧЕТ расписан ия занятий у преподав.	0	437	0	0	0	2991	3633	0	7061	1,21
ПОИСК Учеб – метод. пособий по автору	0	0	336	0	0	0	2010	4081	6427	1,10
РАСЧЕТ колич. опубл. уч. метод пособ препод.	0	0	301	0	0	0	1233	0	1534	0,26
РАСЧЕТ колич опубл. книг препод.	0	0	298	0	0	0	4077	0	4375	0,75
РАСЧЕТ оплаты студ. за год/семес тр	40 9	1076	0	0	0	2578	3340	3045 6	3785 9	6,51
ПОИСК статей по направле нию	0	0	0	563	88 6	0	0	3818	5267	0,90

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru)

2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

исслед.										
РАСЧЕТ колич. опубл. статей препод.	0	0	418	0	0	0	1560	0		
									1978	0,34
РАСЧЕТ колич. опубл. книг препод.	0	0	277	0	0	0	4099	0		
									4376	0,75
ПОИСК книг по автору	0	0	560	528	16 84	0	1260 5	1506 0	3043 7	
										5,23
ПОИСК книг по теме	0	0	435	330	14 79	0	1307 4	1884 8	3416 6	
										5,87
РАСЧЕТ колич. поступ книг	0	0	0	0	56 6	0	1306	2238		
									4110	0,70
РАСЧЕТ колич. книг у читателей	0	0	0	0	62 2	0	0	0		
									622	0,10
РАСЧЕТ числа читателей по вузу	0	0	0	0	47 0	0	0	0		
									470	0,08
РАСЧЕТ успеваем ости в группе	0	230	0	0	0	669	4683	3900 6		
									4458 8	7,66
РАСЧЕТ посещаем ости в группе	0	314	0	0	0	891	5879	3828 3		
									4536 7	7,80
Всего по группам	61 32	1142 3	1120 1	9667	72 26	3514 2	1068 30	3937 61		
Общее									5813 82	100

Таблица 2. Объем транзакции (I, бит)

№	1	2	3	4	5	6	7
I	1060	1080	1070	1060	1080	1070	1010
8	9	10	11	12	13	14	15

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru)

7

2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

1040	980	980	1170	1680	2560	2240	2080
№	16	17	18	19	20	21	22
I	2980	2720	1080	1010	1040	1030	1020
23	24	25	26	27	28	29	30
1080	1060	1060	1020	1040	1020	1020	1280

Таким образом, транзакция каждого типа характеризует частоту появления (статистическую вероятность), количество информации $t_j = (p_j, I_j)$, $j=1...N$ и общий поток (объем)

$$I^* = \sum_{j=1}^N p_j I_j \text{ информации, которую должна обрабатывать}$$

система управления в единицу времени. Задача баланса между I^* и M - мощностью системы управления – рассмотрена в работе / ? /.

Время на обработку одной стандартной элементарной транзакции пропорционально его сложности или количеству информации; при этом считается что информация различных сегментов определена релевантно, то есть равноценна $\tau_j = \tau_0 I_j$.

Причем коэффициент пропорциональности предполагается не зависящим от типа транзакции (что добивается настройкой ИС и рабочего места исполнителя).

Пусть определено K исполнителей (подразделений или пользователей). Обозначим через Q_k – множество операций, выполняемых k -м исполнителем; где $Q = \{Q_k\}$ – распределение операций по исполнителям. Будем называть допустимым такое разбиение, при котором удовлетворяется условие: $\sum_{j \in Q_k} p_j I_j \leq c$, которое показывает что, отдел может обработать ограниченный объем информации в единицу времени. Кроме того, учтем что некоторые транзакции априори могут быть привязаны к конкретным пользователям; т.е.:

$t_{jk} \in Q_k, k = 1, 2, \dots, K, j \in \{1, \dots, N\}$ – это правило определяется вопросами конфиденциальности информации (при этом, только конкретные пользователи авторизованы для работы с конкретными видами информации).

С учетом введенных обозначений дадим формальную постановку задачи:

Решение задачи 1

Определить такое допустимое распределение транзакции $Q = \{Q_k\}$ по исполнителям, что удовлетворяется условие $\sum_{j \in Q_k} p_j \cdot I_j \leq c$ (1) и обеспечивает наиболее равномерную загрузку подразделений; т.е. минимизирует функцию $\Phi = \frac{1}{K} \sum_{k=1}^K \left(\frac{I^*}{K} - \sum_{j \in Q_k} p_j I_j \right)^2 \rightarrow \min$ (2).

Данная задача относится к задаче комбинаторного программирования [5, 6, 7].

А л г о р и т м р е ш е н и я с п о м о щ ь ю П О E x c e l :

1. Нахождение объемов информации (потоков)

1.1. Содержание потока сообщений в месяц определяется как сумма произведений **P** (из табл. 1) на **I** (из табл.2) и приведено в таблице 3:

Таблица 3. Объем информации (поток) за 1 месяц.

	Рек То рат	Уч от дел	Ме тод От дел	На уч от дел	Библио тека	Де ка нат	Пре пода ват (10 фак)	Сту Ден ты (10 тыс)
1	543780	141404 0	433540	445200	2321 40	303266 0	516856 0	426713 60
2	497880	130788 0	333720	338040	0	281124 0	506520 0	348861 60
3	409810	340260	304950	295320	0	297139 0	417514 0	408568 80
4	556500	578760	244648 0	234366 0	0	248464 0	843124 0	407199 00
5	345600	532440	201852	175392	0	214920	768096	389534

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru) 9

2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

			0	0		0	0	40
6	388410	446190	209720 0	201481 0	0	235721 0	707484 0	368219 10
7	107868 0	440360	392890	414100	3797 60	571660	101303 0	0
8	11471 20	60320 0	4222 40	43160 0	405 600	59488 0	10628 80	0
9	49098 0	37240 0	3047 80	28910 0	245 980	23912 0	0	0
10	47628 0	55468 0	3234 00	39102 0	277 340	22834 0	0	0
11	0	42705 0	0	0	0	21785 40	24172 20	10893 870
12	0	80976 0	0	0	0	26997 60	67048 80	21876 960
13	0	22656 00	0	0	0	59750 40	68300 80	0
14	0	16620 80	0	0	0	77436 80	76944 00	0
15	0	12708 80	0	0	0	4908 800	0	0
16	0	13022 60	0	0	0	891 318 0	10826 340	0
17	0	0	91392 0	0	0	0	54672 00	11100 320
18	0	0	32508 0	0	0	0	13316 40	0
19	0	0	30098 0	0	0	0	41177 70	0
20	42536 0	11190 40	0	0	0	268 112 0	34736 00	31674 240
21	0	0	0	57989 0	9125 80	0	0	39325 40
22	0	0	42636 0	0	0	0	15912 00	0
23	0	0	29916	0	0	0	44269	0

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru) 10
2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)

3			0				20	
2			59360	55968	1785		13361	15963
4	0	0	0	0	040	0	300	600
2			46110	34980	1567		13858	19978
5	0	0	0	0	740	0	440	880
2					5773		13321	22827
6	0	0	0	0	20	0	20	60
27	0	0	0	0	6468	0	0	0
2					80			
8	0	0	0	0	4794	0	0	0
					00			
2		23460				682	47766	39786
9	0	0	0	0	0	380	60	120
3						114		
0		40192				048	75251	49002
	0	0	0	0	0	0	20	240
Всег								
о	636040	160834	123979	102061	75097	5436	135406	441401
в гр.	0	00	20	40	80	3320	740,00	180
Общ								683728
ее								880

Постановка задачи 2.

Еще одним критерием распределения операции по рабочим местам является обеспечение требуемой скорости обработки информации с учетом функциональных или технологических связей между операциями. Если в каждой группе Q_k транзакции, относящиеся к одному и тому же бизнес-процессу будут обрабатываться не дольше чем ? с., то общее время на обработку будет cK и вся система будет работать в сбалансированном режиме. Пусть время на обработку одной транзакции будет - t_i , а затраты на сокращение этой длительности на величину τ_i составляет $\varphi_i(\tau_i)$. Тогда задача состоит в отыскании такого разбиения транзакции по группам и

величин τ_i , которое обеспечивало бы требуемый такт с минимальными затратами.

Как описано в [3,4] все транзакции относятся к тем или иным БП, которые в свою очередь можно представить в виде графа, вершины которого соответствуют операциям, а дуги - последовательности выполнения. Другими словами, приступить к обработке очередной транзакции после того, как завершена предшествующая. Как и раньше, будем обозначать через $Q=\{Q_k\}$ - распределение транзакции по группам, удовлетворяющим требуемой последовательности операции.

Решение задачи 2

Определить допустимое распределение транзакции по группам $Q=\{Q_k\}$ и величины τ_i , сокращения продолжительности операции минимизирующее

$$\sum_{i=1}^N \varphi_i(\tau_i) \rightarrow \min \text{ при ограничениях } 0 \leq \tau_i \leq t_i \quad i=1, \dots, N$$

$$\sum_{i \in Q_k} (t_i - \tau_i) \leq c \quad k=1, \dots, K$$

Эта задача о балансе линий конвейера была решена в [3,4]. Воспользуемся ее результатом.

На первом этапе решения задачи произведем декомпозицию графа. Разобьем множество вершин X графа на две непересекающиеся подмножества X_1 и X_2 таким образом, что не существует дуг, идущих из вершин множества X_2 в вершины множества X_1 . Примем что операция $i \in X_1$ выполняется в первых k группах, а операции множества X_2 – в последующих группах. Оценим затраты на обеспечение требуемого такта обработки информации при таком разбиении.

Для этого решим две задачи типа: $\sum_{i \in Y} \varphi_i(\tau_i) \rightarrow \min$ при ограничениях $0 \leq \tau_i \leq t_i, i=1, \dots, N, i \in Y, \sum_{i \in Q_k} (t_i - \tau_i) \leq qc$.

В первой задаче $Y = X_1, q = k$; а во второй $Y = X_2, q = K - k$. Обозначим $\Phi(Y, q)$ -значение в оптимальном решении задачи. Задача при выпуклых функциях $\varphi_i(\tau_i)$ является задачей выпуклого программирования и может быть решена стандартными методами.

Величина $\Phi(X_1, k) + \Phi(X_2, K - k)$ при является оценкой снизу минимальных затрат на сокращение такта до величины c .

На втором этапе решения применим процедуру ветвей и границ с локальной оптимизацией. Пусть (X_1, X_2) – некоторое начальное разбиение. Выделим множество $Q(X_1) \in X_1$, не имеющих исходящих дуг в вершины множества X_1 и множество вершин $Q(X_2) \in X_2$ не имеющих заходящих дуг из вершин множества X_2 . Очевидно, что любое разбиение вида $(X_1 \cup i, X_2 \setminus i)$ допустимо. Содержательно это соответствует допустимому перемещению одной из операций выполняемых на последних K рабочих местах на первые K рабочие мест или наоборот.

Таким образом, из исходного разбиения (X_1, X_2) можно получить $|Q(X_1)| + |Q(X_2)|$ соседних разбиений. Процедура локальной оптимизации заключается в оценке величин каждого разбиения имеющего минимальную величину оценки. Разбиение (X_1, X_2) называется локально-оптимальным разбиением (X_1^*, X_2^*) , $k^*(X_1^*)$, если все соседние с ним имеют не меньшую величину оценки.

Далее, задача рассматривается отдельно для операции множества X_1^* выполняемых на первых $k^*(X_1^*)$ рабочих местах и отдельно для операции X_2^* выполняемых на последних $K - k^*(X_1^*)$ рабочих местах. Алгоритм заканчивается, когда определены множества операции выполняемых на каждом рабочем месте. И допустимая очередность их выполнения на каждом рабочем месте. Дальнейшее улучшение полученного решения можно получить, применяя метод динамического программирования при фиксированной последовательности выполнения операции.

Литература

1. Концепция стратегического развития Республики Казахстан «Казахстан – 2030»// – г.Астана, 1998.
2. Стратегический план развития Международного казахско- турецкого Университета им. А. Ясави на 2009-2013 г. г.// г.Туркестан, 2010.
3. НОВИКОВ Д. А., ЦВЕТКОВ А. В., ГЛАМАЗДИН Е. С. *Управление корпоративными программами: информационные системы и математическое модели*// -М.: Компания Спутник, 2003.-159с.
4. *Технологии и системы информационного корпоративного управления.* 2002. -200 с.
5. ТУКУБАЕВ З. Б., УМАРОВ А. А. *Исследование потоков заявок, поступающих на веб-сервер организации* // Сб. научных трудов “Вестник МКТУ”, №3, с.42-45, г.Туркестан, 2009
6. УМАРОВ А. А., И ДР. *Разработка веб-узла фирмы “АББА* // V Всероссийская школа-семинар молодых ученых “Управление большими системами”. Сб. трудов.-Т.2 Липецк, 2008.

7. ТУКУБАЕВ З. Б., УМАРОВ А. А. *Моделирование и исследование алгоритмов динамического управления потоками сообщений в информационно-вычислительных сетях* // X-я международная научно-техническая конференция «Кибернетика и высокие технологии XXI века». Воронеж. 2009

OPTIMIZATION OF DISTRIBUTION OF RESOURCES AND SPEED OF DATA PROCESSING

Amantur Amangekdievich Umarov, on the example international kazahsko-turkish university of A. Yasawi, Turkestan, Kazakhstan (Unix77@yandex.ru).

Zukirkhan Beisekovich Tukubayev, on the example international kazahsko-turkish university of A. Yasawi, Turkestan, Kazakhstan (zuhr@pochta.ru).

Abstract: In article it is considered problems of optimisation of distribution of streams on groups of users and speed of processing of the information. The algorithm of the decision of a problem with the help on Excel with use of a method of dynamic programming is resulted.

Keywords: Organizational systems, corporate information systems, the consulting companies, the analysis of business processes of the company, information systems of a corporate governance.

1. Амантур Амангелдыевич Умаров, магистр(Unix77@yandex.ru) 15
2. Зухирхан Бейсекович Тукубаев, кандидат технических наук
(zuhr@pochta.ru)