

УДК 681.324

Е.Н. Машенко, доцент, канд. техн. наук**В.И. Шевченко, ассистент***Севастопольский национальный технический университет**ул. Университетская 33, г. Севастополь, Украина, 99053**E-mail: kvf@sevgtu.sebastopol.ua***ОПТИМИЗАЦИЯ УПРАВЛЕНИЯ РАСПРЕДЕЛЕНИЕМ ФАЙЛОВ В
КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СЕТИ**

Рассмотрена модель организации эффективного управления обработкой информационных потоков за счет динамического перераспределения файлов данных в корпоративной сети предприятия. Предложена программная реализация модели.

Ключевые слова: корпоративная сеть, информационный поток, информационная система, эффективное управление.

Современные экономические условия требуют формирования новых подходов к процессам управления предприятиями. Одним из таких подходов является реинжиниринг бизнес-процессов на основе современных информационных технологий. Реализация принципов реинжиниринга бизнес-процессов влечет за собой совершенствование системы информационных потоков, изменение алгоритмов прохождения документов, и следовательно, реинжиниринг информационных систем (ИС), входящих в корпоративную информационную сеть (КИС) предприятия [1]. В этой связи актуальной является задача разработки моделей, позволяющих организовать эффективное управление информационным обменом в КИС с учетом динамически меняющихся характеристик информационных потоков (ИП).

Для решения этой задачи может быть применен агрегативно-декомпозиционный подход, рассмотренный в [2]. В рамках этого подхода информационный обмен в КИС включает набор функциональных задач (ФЗ), процесс решения которых может быть представлен множеством альтернативных графов G . В этом случае задача управления распределением информационных потоков заключается в выборе альтернативы $G^* \subset G$, наиболее эффективной с точки зрения выбранных критериев качества. Схема принятия решения показана на рисунке 1.

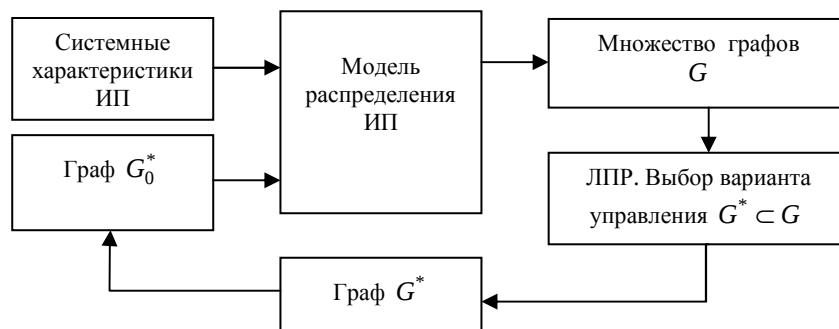


Рисунок 1 – Схема принятия решения по эффективному управлению ИП

Для территориально распределенных организаций на логическом уровне информационный обмен в КИС реализуется как поток запросов к распределенной базе данных предприятия (РБД) и поток данных выборки информации по запросам. Одним из методов повышения эффективности обработки информации в базах данных является оптимальное распределение информационных ресурсов по узлам КИС и организация оптимального доступа к информации в каждом узле. В рамках этого метода может быть рассмотрена задача эффективного распределения файлов данных в КИС произвольной топологии. В работах [2, 3] рассмотрены модели распределения файлов по узлам вычислительной сети, однако в данных моделях сделано предположение о том, что информационный запрос, поступающий в КИС, предполагает обращение только к одному файлу РБД. Как правило, большинство ФЗ в КИС требует обработки информационных потоков, связанных с несколькими файлами данных. Также не учитывается тот факт, что при организации КИС, включающей ряд ИС одинаковой логической структуры, при размещении файлов может быть применен следующий подход: с целью сокращения операций по обработке информационных потоков, которые являются идентичными по предприятию, часть файлов в ИС может быть обобщена и физически размещена на каком-либо из вычислительных узлов хранения данных. Для корректной обработки данных к этим файлам организуется общий доступ. Число и тематика

таблиць файлів даних загального доступу визначаються на етапі реінжиніринга логічної структури КИС. Як правило, це таблиці електронних справочників, що містять інформацію про якісні характеристики номенклатури, про клієнтів бізнес-процесів і ще ряд допоміжної інформації по функціональним задачам підприємства. Частичне об'єднання таблиць даних ІС має як свої переваги, так і недоліки. С однієї сторони, скорочується час введення даних загального користування, з іншої – збільшується час обробки запиту до віддаленої частини ІС. Ще однією особливістю інформаційного обміну в КИС є зміна характеристик інформаційних потоків з часом, що викликає необхідність у динамічному перерозподілі файлів даних.

В роботі пропонується модель розміщення файлів по вузлах КИС, що враховує вищеперелічені особливості інформаційного обміну. Для опису параметрів моделі введені наступні позначення:

$n = 1, N$ – кількість інформаційних систем, що входять до корпоративної мережі підприємства і беруть участь в інформаційному обміні. Припускаємо, що в цілях спрощення синхронізації в РБД, всі ІС в системі мають однакову структуру метаданих;

M – кількість файлів, пов'язаних з інформаційним обміном в екземплярі ІС;

S – кількість типів функціональних задач, що вирішуються в процесі інформаційного обміну;

$k = 1, K$ – кількість автоматизованих робочих місць користувача (АРМ) в КИС;

A_k – k -й АРМ в структурі КИС;

f_{mn} – m -й файл n -ої ІС, пов'язаний з інформаційним обміном в КИС;

$d_{msn}^{(1)}$ – елемент матриці смежності $D_n^{(1)}$, для n -ої ІС, що відповідає графу логічної зв'язності m -го файлу з функціональною задачею S -го типу;

$$d_{msn}^{(1)} = \begin{cases} 1, & \text{якщо } f_{mn} \text{ пов'язаний з функціональною задачею } S \text{ типу} \\ 0, & \text{інакше} \end{cases}$$

$d_{mnk}^{(2)}$ – елемент матриці смежності $D_n^{(2)}$, для n -ої ІС, що відповідає графу розміщення m -го файлу в A_k : $d_{mnk}^{(2)} = \begin{cases} 1, & \text{якщо } f_{mn} \text{ розміщений в } A_k \\ 0, & \text{інакше} \end{cases}$

$D_n^{(0-2)}$ – матриця смежності, що відповідає графу розміщення m -го файлу в A_k в момент часу t_0 ;

$dd_{mnk}^{(2)}$ – елемент матриці $DD_n^{(2)}$, допустимості розміщення m -го файлу в A_k для n -ої ІС:

$$dd_{mnk}^{(2)} = \begin{cases} 1, & \text{якщо } f_{mn} \text{ може розміщуватися в } A_k \\ 0, & \text{інакше} \end{cases}$$

$l = 1, L$ – кількість вузлів зберігання даних в КИС;

$d_{mnl}^{(3)}$ – елемент матриці смежності $D_n^{(3)}$, для n -ої ІС, що відповідає графу розміщення файлу в l -ому вузлі зберігання даних: $d_{mnl}^{(3)} = \begin{cases} 1, & \text{якщо } f_{mn} \text{ розміщений в } l \\ 0, & \text{інакше} \end{cases}$

$D_n^{(0-3)}$ – матриця смежності, що відповідає графу розміщення m -го файлу в l -ому вузлі зберігання даних в момент часу t_0 ;

$dd_{mnl}^{(3)}$ – елемент матриці $DD_n^{(3)}$, допустимості розміщення m -го файлу в файлу в l -ому вузлі зберігання даних для n -ої ІС: $dd_{mnl}^{(3)} = \begin{cases} 1, & \text{якщо } f_{mn} \text{ може розміщуватися в } l \\ 0, & \text{інакше} \end{cases}$

λ_{sk} – інтенсивність надходження запитів, на виконання функціональних задач S -го типу з АРМ A_k ;

u_{snk} – обсяг даних, оброблюваних при виконанні функціональної задачі S -го типу в n -ій ІС в A_k ;

t_{sk}^r – час генерації запиту на виконання функціональної задачі S -го типу в A_k ;

t_{sk}^O – время обработки запроса на выполнение функциональной задачи S -го типа к файлам от A_k ;

t_{sk}^{Π} – время пересылки запроса на выполнение функциональной задачи S -го типа к файлам от A_k ;

T_{mnk}^S – общее время, необходимое для выполнения ФЗ S -го типа в A_k .

T_S^{Π} – пороговое время, выделяемое на выполнения функциональной задачи S -го типа в КИС.

$\bar{V}$ – вектор критериев качества управления информационным обменом.

С учетом общей структуры ИС и возможности совместного доступа к файлам данных справедливо ограничение, показывающее, что в КИС должен присутствовать хотя бы один экземпляр файла данных m -го типа, размещаемый либо в АРМ, либо на одном из узлов обработки данных:

$$\sum_{n=1}^N \sum_{k=1}^K d_{mnk}^{(2)} + \sum_{n=1}^N \sum_{l=1}^L d_{mnl}^{(3)} \geq 1, \quad \forall m = 1, M. \quad (1)$$

Общее время T_{mnk}^S , необходимое для обработки информации, связанной с выполнением функциональной задачи S -го типа в КИС определяется как

$$T_{mnk}^S = t_{smnk}^{\Gamma} + t_{smnk}^O + t_{smnk}^{\Pi} \quad (2)$$

при условии:

$$T_{mnk}^S \leq T_S^{\Pi}. \quad (3)$$

В случае, если условие (3) не выполняется, функциональная задача считается не выполненной и изменения в файлах данных не фиксируются.

Постановка задачи. Объектом исследования является набор из N информационных систем, имеющих одинаковую структуру метаданных, размещенных в КИС произвольной топологии. КИС включает автоматизированные рабочие места пользователя и узлы хранения данных. Файлы информационных систем могут располагаться как в узлах АРМ, так и в узлах хранения данных, возможные варианты размещения определены на этапе синтеза РБД и информация об этих вариантах хранится в матрицах $DD_n^{(2)}$, $DD_n^{(3)}$. На момент времени t_0 файлы размещены в узлах КИС.

Необходимо выбрать вариант размещения файлов данных в КИС, удовлетворяющий условиям (1), (3) и позволяющий организовать эффективный информационный обмен по заданным критериям качества $\bar{V}$, в условиях динамического изменения значений характеристик информационных потоков λ_{sk} и u_{snk} .

Учет динамики функционирования КИС приводит к необходимости использования имитационной модели для принятия решения об эффективном размещении файлов данных. Применение такой модели на этапе реинжиниринга позволит лицу, принимающему решение (ЛПР), имея данные о начальном размещении файлов $D_n^{(0-2)}$, $D_n^{(0-3)}$, данные о возможных вариантах размещения $DD_n^{(2)}$, $DD_n^{(3)}$ и текущие значения характеристик информационных потоков выбрать наиболее эффективный, с точки зрения заданных критериев качества вариант размещения файлов.

Для иллюстрации выбранного подхода, рассмотрим работу имитационной модели на следующем примере: функциональные задачи, связанные с обработкой информационных потоков, условно разбиты на 3 типа:

- 1) задачи, связанные с вводом и корректировкой справочной информации;
- 2) задачи, связанные с выполнением текущих функций предприятия;
- 3) задачи, связанные с выборкой больших объемов аналитической информации.

Наименования файлов РБД КИС и матрица смежности приведены в таблице 1, варианты возможного распределения файлов по АРМ приведены в таблице 2. Предполагается, что КИС состоит из двух АРМ и одного узла хранения данных. Вектора размещения файлов для обоих АРМ идентичны. Основные параметры модели приведены в таблице 3.

Необходимо выбрать наиболее эффективный вариант размещения файлов по критериям максимума среднего объема обрабатываемых в КИС информационных потоков и минимума среднего времени обработки функциональных задач.

Результаты экспериментов приведены в таблице 4. Для рассмотренного примера вариант 5 (см. таблицу 4) является наиболее эффективным. При заданной конфигурации РБД КИС, это будет вариант распределения, при котором все файлы данных размещаются в АРМ, без использования файлов общего доступа в узлах обработки данных.

Таблица 1 – Наименования файлов РБД КИС и матрица смежности

№ файла	Наименование	Связь с функциональной задачей		
		1 типа	2 типа	3 типа
1	Справочник Номенклатура	1	1	1
2	Справочник Единицы измерения	1	1	1
3	Справочник Контрагенты	1	1	1
4	Документ Приходная накладная	0	1	1
5	Документ Расходная накладная	0	1	1
6	Документ Движение денег	0	1	1
7	Отчет Взаиморасчеты	0	0	1
8	Отчет Остатки товаров	0	0	1
9	Отчет Движение товаров	0	0	1
10	Справочник Наши реквизиты	1	1	1
Объем данных функциональной задачи, условных единиц		1...4	1...7	1...10

Таблица 2 – Варианты размещения файлов по АРМ

№ варианта размещения	№ файла									
	1	2	3	4	5	6	7	8	9	10
1	0	0	0	1	1	1	1	1	1	0
2	0	0	1	1	1	1	1	1	1	1
3	0	0	0	0	0	1	1	1	1	0
4	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1

Таблица 3 – Параметры модели

№	Параметр	Значение
1	Единица времени	секунда
2	Время одного прогона	8 часов (28800с)
3	Число ИС	2
4	Количество типов обрабатываемых ФЗ	4
5	Число файлов в базе данных	10
6	Интервал времени между запросами на выполнение ФЗ 1, 2, 3 типа	60 мин, 15 мин, 240 мин экспоненциальное распределение
7	Время генерации запроса на выполнение ФЗ 1, 2, 3 типа в АРМ	Равномерное распределение в интервале от 2 до 10 с
8	Время обработки запроса на выполнение ФЗ 1, 2, 3 типа в АРМ, на условную единицу объема информации	10 с, экспоненциальное распределение
9	Время обработки запроса на выполнение ФЗ 1, 2, 3 типа в узле хранения данных, на условную единицу объема информации	20 с, экспоненциальное распределение
10	Предельное время обработки запроса	1200 с
11	Начальное значение вектора размещения файлов в АРМ	Вариант 1
12	Вектор размещения файлов в узле хранения данных	[1 1 1 0 0 0 0 0 1]

Таблица 4 – Результаты экспериментов

№ варианта размещения	Критерии эффективности			
	Время обработки ФЗ, секунд		Объем данных, условных единиц	
	Среднее	Отклонение	Средний	Отклонение
1	102,883	24,381	3,151	0,412
2	90,761	22,142	3,109	0,414
3	204,386	55,023	3,559	0,354
4	272,127	58,553	3,531	0,396
5	40,250	7,580	3,141	0,325

Анализ полученных результатов позволяет сделать следующие выводы:

1) предлагаемый подход можно использовать для принятия решения по выбору эффективного с точки зрения лица принимающего решение варианта размещения файлов в узлах КИС по заданным критериям эффективности;

2) предлагаемая модель эффективна при решении задач анализа и оптимизации функционирования реальных корпоративных информационных сетей. Для ее использования необходимо собрать достаточный объем статистических данных об обрабатываемых в КИС информационных потоках.

Таким образом, формализована постановка задачи эффективного размещения файлов в узлах КИС при двух критериях, с учетом использования нескольких файлов в одной функциональной задаче и организации общего доступа к файлам информационных систем. Предложена программная реализация имитационной модели и проведены эксперименты для модели информационного обмена в РБД КИС упрощенной структуры.

Перспективы дальнейших исследований. Предполагается расширить комплекс моделей управления распределением информационных потоков для различных стратегий управления и структур КИС в условиях динамически меняющихся характеристик информационных потоков. На основе полученного комплекса моделей предполагается создание системы поддержки принятия решений, по эффективному управлению распределением и обработкой информационных потоков в КИС.

Библиографический список

1. Ахтырченко К.В. Методы и технологии реинжиниринга ИС / К.В. Ахтырченко, Т.П. Сорокваша // Труды Института системного программирования. — М, 2003. — С. 116–128.
2. Цвиркун А.Д. Имитационное моделирование в задачах синтеза структуры сложных систем (оптимизационно-имитационный подход) / А.Д. Цвиркун, В.К. Акинфеев, В.А. Филиппов. — М: Наука, 1995. — 173 с.
3. Цегелик Г.Г. Системы распределенных баз данных / Г.Г. Цегелик. — Львов: Свит, 1990. — 168 с.

Поступила в редакцию 21.10.2009 г.