

ВОПРОСЫ ПОСТРОЕНИЯ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННЫХ СИСТЕМ НА ОСНОВЕ ПЕРЕДВИЖНЫХ ПУНКТОВ УПРАВЛЕНИЯ

М.А. СОНЬКИН, В.К. ПОГРЕБНОЙ

Национальный исследовательский
Томский политехнический университет
sonkin@tpu.ru

Территориально-распределенные системы, включающие не только стационарные, но и передвижные объекты, получают все большее распространение. На примере представления иерархической структуры системы в виде графовой модели и условий баланса выделяемых и потребляемых коммуникационных ресурсов предложено два алгоритма для проектирования архитектуры разрабатываемых систем. Предложенные методы и реализующие их алгоритмы эффективно работают с размерностями реальных задач и учитывают плотность расположения объектов разного типа на топологическом поле.

Ключевые слова: распределённые системы управления, проектирование архитектуры системы, оптимизация передачи данных.

Развитие беспроводных сетей и миниатюризация аппаратуры создали новые технические и технологические возможности для построения навигационно-телекоммуникационных комплексов (НТК) нового поколения (НП). Составной частью НТК НП могут быть, например, аппаратно-программные средства мониторинга и управления мобильными комплексами (МК), передвижные пункты управления (ППУ) и беспилотные летательные аппараты (БПЛА). Особым спросом НТК пользуются в орга-

низациях и системах быстрого реагирования на природные, техногенные и иные критические ситуации.

Для повышения функциональных возможностей по управлению и координации действий нескольких ППУ и МК, находящихся в одном районе, а также применения БПЛА для мониторинга территории с воздуха целесообразно создание соответствующей архитектуры НТК НП.

Обеспечение эффективности управления в рамках НТК НП может быть достигнуто путем использования комбинированного способа взаимодействия: сочетания стратифицированного управления с горизонтальным взаимодействием между уровнями (стратами). На рис. 1 приведен пример обобщенной структуры НТК НП с комбинированным способом управления. В соответствии с этим необходимо обеспечить:

- синхронизацию режимов приема-передачи и обработки навигационной, фото- и видео- информации в режиме реального времени на уровне наземной станции управления БПЛА, управляющей станции мобильного комплекса, диспетчерского центра;
- передачу управления БПЛА от одного диспетчерского центра другому;
- интеграцию каналов связи различного типа.

Совершенно очевидно, что система, приведенная на рис. 1, относится к классу территориально-распределенных систем (ТРС), а в качестве ее отличительных свойств можно выделить принципы иерархичности, централизма и многосвязности. Данные свойства во многом предопределяют концептуальную основу построения архитектуры рассматриваемой

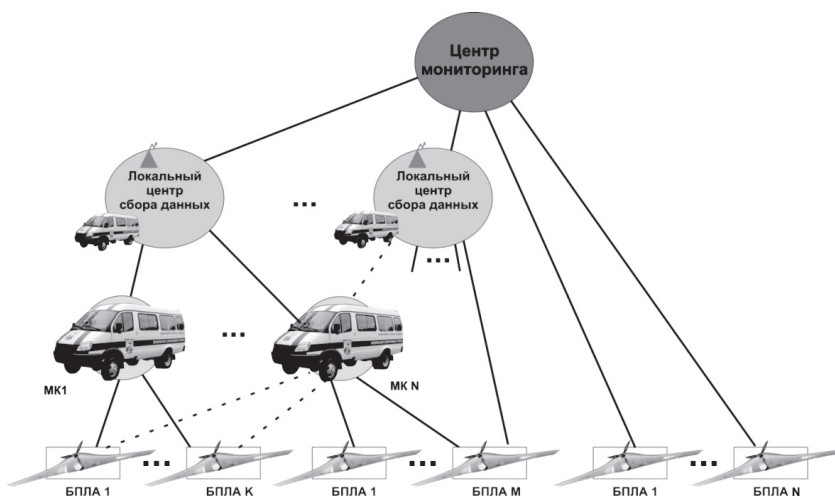


Рис. 1. Пример обобщенной структуры НТК НП с комбинированным способом управления

го класса ТРС. Вместе с тем свойства иерархичности и многосвязности предполагают наличие альтернативного выбора многих архитектурных решений при построении конкретной системы.

При построении и оптимизации архитектуры ТРС важное место занимает задача разбиения множества объектов нижнего уровня иерархии (**терминальных объектов** – МК) на подмножества с целью подключения объектов этих подмножеств к объектам более высокого уровня (**центрам** – ППУ). Последовательно решая эту задачу относительно объектов на более высоком уровне, мы получим новые центры и, в конечном итоге, – иерархическую структуру системы. Центр и подключенное к нему подмножество объектов более низкого уровня в последующем будем именовать **локальным узлом иерархической структуры**. Определение состава локальных узлов в иерархической структуре в общем случае связано с задачей декомпозиции множества объектов на подмножества. На первом этапе множество объектов составляют терминальные объекты, а на последующих этапах декомпозиции в качестве множества объектов выступают центры.

В терминах теории графов как объекты (вершины графа), так и отношения между ними (ребра графа) при описании структуры системы сопровождаются указанием ряда атрибутов. Такие атрибуты отражают конкретные свойства соответствующих объектов и связей между ними. Графы с приписанным набором атрибутов будем именовать графовой моделью.

Графовые модели позволяют учесть параметры, характеризующие объекты системы и отношения между ними, и с нужной подробностью описывают архитектуру системы. Если при этом атрибуты, как в нашем случае, будут отражать параметры объектов и каналов связи между ними, то мы вправе проводить параллель между графовой моделью и архитектурой системы.

Описание и анализ объектов и связей между ними будем вести относительно примера визуального представления графовой модели иерархической структуры системы, изображенной на рис. 2.

Объекты структуры ТРС ассоциируются с действиями в системе и, по аналогии с сетями Петри, на рис. 2. изображаются планками. Для удобства объекты (планки) в последующем будем именовать модулями, имея в виду аппаратно-программные средства объекта, которые осуществляют сбор, обработку, анализ, формирование сообщений и другие операции. Предполагается также, что модули являются сложными и могут быть развернуты в графовые модели, которые более детально и функционально-ориентированно описывают работу объекта.

Кружками на рис. 2, также по аналогии с сетями Петри, данные, формируемые терминальным объектом для передачи в центр, представлены

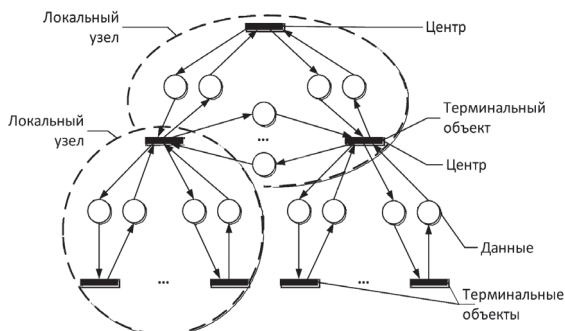


Рис. 2. Пример визуального представления графовой модели

одним кружком (позицией). В случае развернутого представления графовой модели все многообразие данных будет представлено в явном виде отдельными позициями. Модули центров, как показано на рис. 2, формируют несколько данных для передачи вверх и вниз по иерархической структуре, в том числе и для модулей, расположенных на одном уровне. Передача данных между модулями одного уровня, как правило, может осуществляться внутри одного локального узла иерархической структуры.

Для примера приведем параметры, с которыми придется оперировать при построении архитектуры системы. Допустим, это параметры имеющихся ресурсов:

- процессорное время τ_j , которое выделяется для выполнения модуля m_j ;
- объём памяти ρ_j запоминающего устройства j -го объекта;
- каналный ресурс K_{jr} определяет предельный объём данных, который можно передать по r -му каналу из j -го объекта.

Параметры ресурсов устанавливаются как предельные значения для определенного отрезка времени, который будем именовать интервалом анализа τ . Величина τ подбирается так, чтобы циклически повторяющиеся процессы в системе укладывались в интервал анализа целое число раз.

Выполнение каждого модуля инициируется (запускается) по определенным правилам, в соответствии с технологическим регламентом работы системы. Будем различать запуски циклические, вероятностные, детерминированные и условные. Циклические запуски модулей m_j осуществляются через интервалы времени $\Delta\tau_j$. Для вероятностных запусков указываются вероятности наступления определенных событий, диктующих необходимость выполнения модуля. Детерминированные запуски также являются вероятностными, но отличаются наличием дополнительных правил много-

кратного запуска модуля в определенные моменты времени. Условные запуски могут быть отнесены к вероятностным, но они наступают не под воздействием внешних по отношению к системе факторов, а под влиянием внутренних, условия наступления которых контролирует сама система.

Для разных модулей m_j интервалы циклов запуска $\Delta\tau_j$ могут различаться. Например, если цикл запуска одних модулей составляет 10 единиц, а других – 20, то интервал анализа τ может быть принят равным 100 единицам. Исходя из этого для каждого модуля m_j , суммируя все виды запусков, можно подсчитать частоту ρ_j , с которой он будет выполняться в одном интервале анализа. Так, для интервала анализа $\tau = 100$, частота ρ_j срабатывания модуля m_j с $\Delta\tau_j = 10$ будет складываться из десяти циклических и, например, в среднем двух запусков по другим основаниям, т.е. $\rho_j = 10+2=12$.

Каждый запуск модуля m_j будем сопоставлять с формированием состояния, данного по одной из выходных позиций. Поэтому число состояний данного ρ_j формируемого в позиции d_i за один интервал τ , не превысит значение ρ_j . Если модуль m_j формирует состояния для V_j , то можно записать, что $\sum_{d_i \in V_j} \rho_i = \rho_j$.

Значения величин ρ_i и ρ_j дают возможность оценить объемы потребления ресурсов в системе. Для этого введём следующие параметры потребления ресурсов:

τ_{ji} – процессорное время j -го объекта, затрачиваемое для формирования одного состояния в позиции d_i ;

ρ_{ji} – память, необходимая для хранения одного состояния позиции d_i в j -том объекте;

K_{jir} – потребление канального ресурса K_{jr} при передаче по r -каналу одного состояния позиции d_i , сформированного модулем m_j .

Перечисленные параметры позволяют сопоставить наличие баланса между имеющимися ресурсами и их потреблением. Выражения для проверки баланса имеют следующий вид:

$$\alpha_j \sum_{d_i \in V_j} \tau_{ji} \rho_i \leq \tau_j, \quad j = 1, 2, \dots, m; \quad (1)$$

$$\beta_j \sum_{d_i \in V_j} \rho_{ji} \rho_i \leq P_j, \quad j = 1, 2, \dots, m; \quad (2)$$

$$\gamma_j \sum_{d_i \in V_j} K_{jir} \rho_i c_i \leq K_{jr}, \quad j = 1, 2, \dots, m, \quad r = 1, 2, \dots, R. \quad (3)$$

Здесь $\alpha_j, \beta_j, \gamma_j$ – коэффициенты резервирования ресурсов по параметрам τ_j, P_j, K_{jr} ;

c_i – коэффициент селекции, определяющий снижение числа сформированных в позиции d_i состояний при передаче их потребителю;

R – число типов каналов, используемых в системе.

Каждый j -тый объект, помимо канального ресурса K_{jr} по передаче данных, должен обладать канальным ресурсом K'_{jr} по приему данных, посту-

пающих от совокупности V'_j входных позиций $d_i \in V'_j$ и потребляющих ресурс K'_{jr} в объеме K'_{ir} . Условие проверки баланса запишется аналогично (3) в виде

$$\gamma'_j \sum_{d_i \in V'_j} K'_{jr} \rho_i c_i \leq K'_{jr}, \quad j = 1, 2, \dots, m, \quad r = 1, 2, \dots, R. \quad (4)$$

В условии баланса (1) предполагается, что процессорное время в размере τ_{ji} затрачивается для формирования состояния в каждой выходной

позиции $d_i \in V_j$. Если для разных позиций d_i формируется одно состояние, то в этом случае экономится ресурс τ_j , и при необходимости такая ситуация может быть учтена. При разработке конкретной ТРС возможны и другие особенности потребления ресурсов, которые могут учитываться при составлении уравнений баланса.

Проверку условий (1) – (4) можно проводить для объектов существующей системы при ее модернизации или для анализа очередного варианта иерархической структуры проектируемой ТРС. Анализ существующей системы, в том числе и при ее модернизации, позволяет оценить «узкие места» и найти решения по их преодолению. Но если будут разработаны методы принятия решений при проектировании ТРС, то, несомненно, они могут быть применены и для модернизации.

Для реализации сопоставления выделяемого и потребляемого ресурса важно найти эффективные методы, которые на стадии формирования варианта иерархической структуры учитывали бы условия баланса ресурсов. При разработке таких методов будем исходить из того, что ТРС строится на базе априорно заданного числа терминальных объектов. Такие объекты (например, метеостанции) являются интеллектуальными терминалами и осуществляют сбор и обработку первичных данных для передачи в центр. Терминальные объекты также могут принимать данные из центра в виде инструкций, сообщений, запросов и т.п.

В общем случае будем считать, что заданное множество объектов M не является однородным и разбивается на непересекающиеся подмножества M_s , $s=1, 2, \dots, S$ по типам. Основанием для деления терминальных объектов по типам могут быть различия в функциональных возможностях, различия в доступе к определенным видам каналов для передачи и приема данных и другие различия.

Для рассматриваемой задачи важно различать объекты по ресурсным параметрам. Например, функциональные признаки влияют на ресурсы τ

и τ' , P и P' , а возможности доступа к каналам – на ресурсы K и K' . Поэтому в последующем будем считать, что все различия объектов сведены к определенным значениям ресурсных параметров. При этом если объекты различаются по какому-либо признаку, который не влияет на такие параметры, то это различие не учитывается.

Сведения о потреблении ресурсов терминальными объектами, составляющими множество M , удобно представить в виде таблицы. Для наглядности таблица приведена в качестве примера описания ресурсных параметров для $|M|=50$ объектов, разбитых на 5 видов подмножеств M_s . Число объектов в каждом подмножестве M_s обозначено величиной m_s , которая приведена в отдельной колонке в правой части таблицы. Здесь же в правой колонке приведена величина $\rho_s = \sum_{m_j \in M_s} \rho_j$, которая равна суммарной

частоте выполнения модулей m_j , входящих в подмножество объектов M_s , за один интервал анализа.

Обозначения параметров в таблице записаны, исходя из предположения, что каждый терминальный объект имеет одну выходную позицию с номером d_1 и одну входную с номером d_2 (см. рис. 2). При обозначении величины τ_j вместо j указано s -е подмножество, к которому принадлежит j -й объект. Передачу и прием данных объекты могут вести по четырем типам каналов, $R=4$. Наличие в таблице нулевых значений параметров канальных ресурсов означает, что соответствующий тип канала не используется объектами подмножества.

Подводя итог описанию и анализу терминальных объектов, можно отметить, что исходные сведения о выделяемых и потребляемых ими ресурсах определены, и для всех объектов множества M выполняются условия (1) – (4). Согласно иерархической структуре, изображенной на рис. 2, терминальные объекты подключаются к нескольким центрам. Предполагается, что каждый центр способен подключать определенное число терминальных объектов, порождая при этом локальный узел иерархической структуры.

Определение рациональной совокупности локальных узлов, охваты-

Таблица

Потребление ресурсов терминальными объектами

S	τ_{s1}	τ'_{s2}	P_{s1}	P'_{s2}	$K_{1,1}$	$K_{1,2}$	$K_{1,3}$	$K_{1,4}$	$K'_{2,1}$	$K'_{2,2}$	$K'_{2,3}$	$K'_{2,4}$	m_s	ρ_s
1	8	5	15	3	0	3	2	0	0	1	1	0	9	45
2	10	5	12	7	2	0	1	1	1	0	2	1	7	64
3	2	4	23	4	0	1	0	3	0	1	0	2	15	38
4	8	6	14	5	1	2	0	0	2	2	0	0	9	96
5	7	9	18	8	3	0	2	0	1	0	1	0	10	80

вающих все терминальные объекты множества M , по существу является основной целью первой задачи, сформированной в начале главы. Решение такой сложной задачи предполагается разбить на два этапа. На первом этапе определяется минимальное число центров, необходимых для подключения всех объектов множества M . При этом предполагается, что каждый центр имеет ограниченное число объектов. Второй этап связан с разбиением множества M на подмножества, каждое из которых подключено к одному из центров. Решение задач второго этапа рассматривается с учетом территориального расположения объектов множества M и будет изложено позже.

Рассмотрим задачи первого этапа, связанные с определением числа центров. Поиск решения предлагается вести на основе сопоставления ресурсов терминальных объектов и ресурсов центров. Значения ресурсных параметров терминальных объектов известны. Если установить значения ресурсных параметров центров, то нетрудно определить ориентировочное число необходимых центров. Поэтому основное внимание следует сосредоточить на установлении значений ресурсных параметров центров. Заметим, что по составу ресурсные параметры центров ничем не отличаются от терминальных объектов. На основе характеристик производительности аппаратуры, принятой для комплектации центра, частоты срабатывания модуля при формировании состояний данных выходных позиций и обработки входных позиций, объёмов передаваемой и поступающей информации, загружающей каналы ресурсы, можно для принятого интервала анализа рассчитать значения параметров ресурсов, которыми располагает обобщенный центр. Эти параметры обозначим величинами τ^* , P^* , K_r^* , K'_r^* .

Общий объем ресурсов, потребляемых всеми центрами (одним виртуальным центром), вычисляется аналогично тому, как это делалось для терминальных объектов. Тогда число центров, получаемое для каждого вида ресурса, определяется отношением общего объема потребляемого ресурса и имеющегося в обобщенном центре:

$$\alpha(\sum_{d_i \in V} \tau_i \rho_i) / \tau^* = G(\tau); \quad (5)$$

$$\beta(\sum_{d_i \in V} \rho_{ji} \rho_i) / P^* = G(P); \quad (6)$$

$$\gamma(\sum_{d_i \in V} K_{ir} \rho_i) / K_r^* = G(K); \quad (7)$$

$$\gamma'(\sum_{d_i \in V} K'_{ir} \rho_i) / K'^* = G(K'). \quad (8)$$

Здесь V, V' – множества входных и выходных позиций по отношению к объектам множества M ;

$G(\tau), G(P), G(K), G(K')$ – число обобщенных центров, необходимое для обслуживания терминальных объектов множества M ресурсом τ, P, K, K' соответственно;

$\alpha, \beta, \gamma, \gamma'$ – коэффициенты резервирования ресурсов, которые в данном случае должны учитывать возможные выходы и входы в центры, расположенные на этом же уровне иерархии, и в вышестоящие центры, относительно которых рассматриваемые обобщенные центры выступают в роли терминальных объектов.

Значения параметров $\tau_p, P_p, K_{ip}, K_{ip}'$ определяются с использованием параметров входных и выходных позиций терминальных объектов. Так, если значения τ_s определяют затраты процессорного времени для формирования одного состояния позиции d_2 для всех объектов множества M_s , то выражение (5) запишется в виде

$$\gamma \left(\sum_{s=1}^S \tau_s \rho_s \right) / \tau^* = G(\tau).$$

Для $\tau_l = \tau_5 = 0,1; \tau_2 = 0,4; \tau_3 = 0,3; \tau_4 = 0,5$ при $\tau^* = 25, \alpha = 1,4$ и ρ_s из таблицы имеем $1,4(0,1 \cdot 45 + 0,4 \cdot 64 + 0,3 \cdot 38 + 0,5 \cdot 96 + 0,1 \cdot 80) / 25 = 5,46$. Значение $G(\tau) = 5,46$ означает, что 6 центров с ресурсом процессорного времени 25 единиц в одном интервале анализа своевременно справятся с формированием состояний выходных позиций d_2 (входных для объектов множества M). Аналогично по выражениям (6) – (8) определяется число центров относительно других ресурсов. Сопоставляя значения величин $G(\tau), G(P), G(K), G(K')$, принимается решение о величине G , которая определяет число центров.

Рассмотрим ситуацию, когда терминальные объекты настолько отличаются друг от друга, что могут быть подключены только к тому центру, который оснащен соответствующими возможностями. Например, если некоторый объект формирует и передает в центр видеоданные, то нет необходимости все центры оснащать эффективными средствами их обработки. Получается, что центры обладают различными возможностями для подключения объектов. Способность центра g -го вида подключать определенное количество терминальных объектов разного типа представим вектором подключения $A_g = \{a_{gs}\}; g = 1, 2, \dots, G^*; s = 1, 2, \dots, S$. Элемент a_{gs} вектора A_g равен числу объектов s -го вида, которые могут быть подключены к центру g -го вида.

В этом случае для определения числа центров более перспективным представляется другой подход. В его основе лежит предположение о наличии заранее сформированного множества G^* видов центров. Формирование множества G^* может осуществляться на этапе разработки ТРС

виртуально, на основе анализа различия объектов в множестве M . При этом номенклатура виртуальных видов центров может существенно превышать реальную потребность в центрах, которая ассоциируется с числом G . Для генерации множества G^* могут использоваться различные эвристические правила, которые здесь не рассматриваются.

Для ряда предметных областей, в которых проектируется ТРС, множество G^* формируется на промышленной основе, то есть возможные для применения виды центров известны разработчику проекта заранее. В этой ситуации из множества G^* следует выбрать минимальное число центров, способных подключить все объекты множества M .

Введем переменную x_g , $g=1,2,\dots, G^*$, которая определяет число центров g -го вида, выбранных для подключения объектов множества M . Тогда задачу выбора минимального числа центров, необходимых для подключения объектов множества M , можно записать в виде

$$\sum_{g=1}^{G^*} x_g \Rightarrow \min; \quad (9)$$

$$\sum_{g=1}^{G^*} a_{gs} x_g \geq m_s, \quad s = 1, 2, \dots, S; \quad (10)$$

$$x_g - \text{целое положительное для всех } s=1, 2, \dots, S \quad (11)$$

Задача (9)–(11) относится к классу задач целочисленного линейного программирования. В результате решения задачи получим вектор $X^* = \{x_{gs}^*, g=1, 2, \dots, G^*, \text{ компоненты которого } x_g^* \text{ определяют минимальное число центров } g\text{-го вида, необходимых для подключения объектов множества } M.$

Критерий (9) записан при условии, что все центры равнозначны по цене, либо такие различия не существенны. Если это не так, то вводится цена C_g центра g -го вида, и критерий (9) в этом случае запишется в виде

$$\sum_{g=1}^{G^*} C_g x_g \Rightarrow \min.$$

Ограничение (10) допускает некоторую функциональную избыточность центров, когда сумма s -тых компонентов a_{gs} векторов подключения A_g превышает число объектов m_s в множестве M_s . Такая избыточность может рассматриваться в качестве резерва, либо по возможности исключаться при комплектации центров.

Полученная в результате решения задачи (9)–(11) совокупность цен-

тров $G = \sum_{g=1}^{G^*} x_g^*$; при необходимости может рассматриваться в каче-

стве терминальных объектов для получения центров на более вы-

соком уровне иерархической структуры, как это показано на рис. 2. Аналогично в качестве терминальных объектов рассматриваются центры, которые были получены ранее в количестве G штук на основе сопоставления ресурсов. Изложенные выше методы используются и при определении числа центров на более высоком уровне иерархии.

В предыдущем параграфе было определено число центров, способных подключить терминальные объекты проектируемой ТРС. Предполагается также, что каждый центр способен подключить определенное количество объектов. Если все объекты одного типа, каждый центр подключает примерно равное число объектов. В условиях, когда объекты различаются по типам, состав объектов, подключаемых к центру, определяется его вектором подключения. В общих случаях терминальные объекты должны быть разбиты на подмножества, каждое из которых подключается к центру и образует локальный узел иерархической структуры.

Таким образом, задача построения локальных узлов сводится к разбиению объектов на подмножества. Основным фактором, определяющим критерий разбиения, является размещение объектов на определенной территории. Территорию, на которой расположены терминальные объекты, будем именовать топологическим полем. Очевидно, что предпочтительным является такое разбиение объектов на подмножества, при котором центры и подключаемые к ним подмножества располагаются на топологическом поле в компактных областях.

На содержательном уровне задача размещения центров на топологическом поле и подключения к ним объектов формулируется просто. На топологическом поле расположено множество $Q = \{q_{ir}\}$, $i=1, 2, \dots, n$, $r=1, 2, \dots, R$, состоящее из подмножеств Q_r объектов r -го типа. Для объектов $q_{ir} \in Q$ известны координаты (x_i, y_i) их расположения на топологическом поле. Расстояние d_{ij} между объектами q_{ir} и q_{jr} вычисляется по выражению

$$d_{ij} = ((x_i - x_j)^2 + (y_i - y_j)^2)^{\frac{1}{2}}.$$

Размещаемая совокупность центров представлена множеством $C = \{c_s\}$, $s=1, 2, \dots, S$ с векторами подключения $A_s = \{a_{sr}\}$. Здесь S – общее число центров, необходимое для подключения множества объектов Q . Требуется для каждого центра C_s найти такие координаты (x_s, y_s) , чтобы сумма его расстояний до подключаемых объектов была минимальной.

В предлагаемой стратегии поиска решения будем исходить из того, что первичным является размещение центров C_s , т.е. определение для них координат (x_s, y_s) , а после этого решается задача подключения объектов к

этим центрам. Начнем рассмотрение с более простой задачи, когда все объекты в множестве Q одного типа, т.е. $R=1$. Центры C_s в этом случае могут различаться только числом подключаемых объектов.

На начальной стадии поиска варианта размещения центров принимается, что они располагаются в местах размещения объектов, т.е. координаты центров совмещаются с координатами объектов, выбранных для размещения центров. Объекты, в которых располагаются центры, будем именовать полюсами. Число полюсов совпадает с числом центров.

В общем случае любой объект может быть объявлен полюсом. В свою очередь, каждый полюс рассматривается в качестве места для размещения центра и подключения к нему объектов. Поэтому после подключения объектов к полюсам, каждый из них окажется в отдельном подмножестве. Если с этой позиции рассматривать любые два объекта, то чем дальше они удалены друг от друга, тем с большей вероятностью они попадут в разные подмножества, т.е. будут подключены к разным центрам. Это же можно сказать и относительно полюсов: чем дальше они расположены друг от друга, тем более обоснованным оказывается их присутствие в разных подмножествах.

Это соображение хорошо поясняется на подмножествах, представленных на рис. 3. Например, объекты q_1 и q_{14} имеют больше оснований попасть в разные подмножества, чем близко расположенные объекты q_{21} , q_{22} или q_8 , q_9 . Аналогично, если в качестве полюсов выбраны объекты q_1 , q_5 , q_{18} , q_{24} , q_{22} , то они «естественным путем» попадают в разные подмножества, выделенные на рис. 3 пунктирными линиями, а не только потому, что эти объекты являются полюсами. Например, относительно объектов q_1 , q_2 , q_6 , q_{11} , q_{17} , назначенных полюсами произвольным образом, говорить о том, что они «естественным путем» попадут в разные подмножества с хорошими оценками компактности, не приходится.

Исходя из этого, алгоритм формирования полюсов, реализующий данное эвристическое правило, устроен так, что выбирает необходимую совокупность полюсов, в которой они как можно дальше отстоят друг от друга. Алгоритм реализован программой «Полюс» и получил обозначение П1. Формирование множества

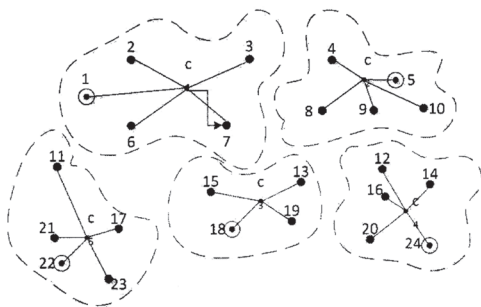


Рис. 3. Пример визуального разбиения объектов на подмножества

полюсов $Q_p = \{q_{ip}\}$, $p=1, 2, \dots, P$ с помощью алгоритма α_i осуществляется последовательно и включает выполнение следующих операций:

Первым полюсом q_{i1} принимается объект $q_i \in Q$, который максимально удален от всех других объектов множества Q . Объект q_{i1} выбирается согласно выражению

$$\forall q_i \in Q [\sum_{q_j \in Q \setminus q_i} d_{ij} = \max] \Rightarrow (q_i - q_{i1}) . \quad (12)$$

Здесь d_{ij} – расстояние между объектами q_i и q_j . Объект q_{i1} заносится в множество полюсов Q_p .

Для каждого объекта $q_i \in Q \setminus Q_p$ определяются расстояния $d_{i,ip}$ от объекта q_i до полюсов $q_p \in Q_p$. Среди этих расстояний выбирается минимальное $d_{i,ip}^*$ и запоминается. Таким образом, каждому объекту $q_i \in Q \setminus Q_p$ соответствует величина $d_{i,ip}^*$.

Среди множества величин $\{d_{i,ip}^*\}$ выбирается максимальное значение, а соответствующий объект $q_i \in Q \setminus Q_p$ принимается в качестве очередного p -го полюса q_{ip} и включается в множество Q_p .

Операции 2 и 3 повторяются до тех пор, пока число полюсов в множестве Q_p не достигнет величины $p = |Q_p|$, то есть станет равным числу размещаемых центров.

Заметим, что при выборе 2-го полюса в множестве Q_p присутствует только один элемент q_{i1} . Поэтому в данном случае «выбор» минимально-

го расстояния $d_{i,ip}^*$ (см. п.2 алгоритма) для объекта $q_i \in Q \setminus Q_p$ сводится к одной величине $d_{i,i1}$. На следующей иерархии алгоритма, при определении 3-го полюса, выбор будет производиться из двух величин $d_{i,i1}$ и $d_{i,ip2}$, и так далее.

Равномерность расположения полюсов по площади поля является основным свойством алгоритма П1. Это свойство можно отнести к достоинству по отношению к полям с регулярной плотностью расположения объектов и к недостатку по отношению к нерегулярному расположению объектов. Алгоритм П1 оказался «нечувствительным» к неравномерной плотности расположения объектов на поле. Поэтому при существенно неравномерной плотности расположения объектов как альтернатива алгоритму П1 разработан алгоритм П2.

Отличие алгоритма П2 от П1 заключается в том, что после выбора по выражению (12) первого полюса выполняется подключение к нему ближайших объектов согласно числу объектов, подключаемых к соответству-

ющему центру C_1 , и полученное подмножество объектов исключается из дальнейшего рассмотрения. Для выбора второго полюса на оставшемся множестве объектов повторно выполняется операция выбора первого полюса по выражению (12), определяется подмножество ближайших объектов для центра C_2 и исключается из рассмотрения. Выбор последующих полюсов и исключение подмножеств объектов производится аналогично.

Для повышения эффективности работы алгоритма необходимо решить задачу подключения объектов к полюсам (центрам) так, чтобы сократить общую протяженность линий связи на подключение.

Считаем, что на топологическом поле с помощью алгоритма П1 или П2 определено множество полюсов $Q_p = \{q_{ip}\}$, $p=1, 2, \dots, P$. Число полюсов P равно числу центров S . Для каждого центра c_s , $s=1, 2, \dots, S$ известно число m_s объектов $q_i \in Q$, $i=1, 2, \dots, n$, которое он может подключить. В частности, число m_s для всех центров может быть одинаковым. Определены также расстояния d_{si} между центрами, закрепленными за полюсами q_{ip} , и объектами q_i .

Введем переменную x_{si} , $x_{si}=1$, если объект q_i подключен к центру c_s , $x_{si}=0$, в противном случае. Задача подключения объектов к центрам с учетом принятых обозначений формулируется как задача математического программирования:

$$\sum_{s=1}^S \sum_{i=1}^n d_{si} x_{si} \Rightarrow \min; \quad (13)$$

$$\sum_{i=1}^n x_{si} = m_s, \quad s = 1, 2, \dots, S; \quad (14)$$

$$\sum_{s=1}^S x_{si} = 1, \quad i = 1, 2, \dots, n. \quad (15)$$

Критерий (13) минимизирует суммарные расстояния от центров до подключаемых к ним объектов. Уравнение (14) обеспечивает подключение к каждому центру c_s ровно a_s объектов. Уравнение (15) обеспечивает подключение каждого объекта q_i только к одному центру.

Задача в постановке (13) – (15) относится к классу легко решаемых задач транспортного типа, например, с помощью венгерского алгоритма. В этой задаче число объектов a_s , подключаемых к центрам, ассоциируется с объемами производства, а объекты соответствуют единичным объемам потребления. Объемы перевозок x_{si} в нашей задаче могут составлять не более одного объекта. Соответственно, переменные x_{si} являются булевыми.

В тех случаях, когда число объектов на топологическом поле измеряется сотнями или даже тысячами, возникает потребность уменьшения размерности задачи без заметного ухудшения качества ее решения. С этой целью на топологическое поле наносится сетка с определенным размером шага так, чтобы число ячеек, содержащих объекты, было существенно меньше общего числа объектов. Таким образом, изменяя шаг ячейки, можно подобрать приемлемую размерность задачи. Каждая непустая ячейка объявляется новым объединенным объектом с координатами центральной точки ячейки.

Задача (13) – (15) решается относительно новых пунктов потребления, представленных объединенными объектами (непустыми ячейками). Переменная x_{si} в этом случае определяет число объектов, взятых из i -й ячейки для подключения к центру c_s . В правой части уравнения (15) вместо единицы указывается число объектов в ячейке. Задача (13) – (15) в такой постановке полностью соответствует классической задаче математического программирования транспортного типа.

Разбиение объектов Q на подмножества Q_s , полученное в результате решения задачи (13) – (15), зависит от расположения полюсов. Как правило, полюса в подмножествах Q_s относительно критерия (13) располагаются не лучшим образом. Это хорошо видно на рис. 2, где подмножества сформированы на основе полюсов $(q_{1,1}, q_{5,2}, q_{18,3}, q_{24,4}, q_{22,5})$. Например, в подмножестве Q_2 для размещения центра вместо объекта q_5 лучше выбрать объект q_9 . Если имеется возможность отказаться от размещения центра c_s в одном из объектов подмножества Q_s , то для этого подмножества определяется центральная точка q_s^* , как это показано на рис. 3. Координаты (x_s, y_s) точки q_s^* вычисляются для каждого подмножества Q_s как центры при условии, что масса объектов $q_i \in Q_s$ принимается равной единице, т.е.

$$x_s = 1/m_s \sum_{q_i \in Q_s} x_i; \quad y_s = 1/m_s \sum_{q_i \in Q_s} y_i; \quad m_s = |Q_s|. \quad (16)$$

Заметим, что координаты (x_s, y_s) , полученные по (16), не соответствуют строго центральной точке q_s^* для объектов множества Q_s . В отличие от

точки q_s^* для точки q_s величина $\sum_{q_i \in Q_s} d_{si}$ не является минимальной. Здесь d_{si} – расстояние между объектом $q_i \in Q_s$ и точкой q_s .

Анализируя приведенные выше рассуждения относительно подмножеств Q_s , как локальных узлов иерархической структуры, напрашивается

следующий вывод. Возможна и другая стратегия поиска решений данной задачи. В этой стратегии первичным является определение компактных подмножеств Q_s , а размещение в каждом из них центров легко осуществляется путем выбора одного из объектов подмножества или на основе выражения (16).

Таким образом, согласно стратегии, рассматриваемая задача заключается в следующем. Множество объектов Q необходимо разбить на совокупность подмножеств $\{Q_s\}$, $s=1,2,\dots,S$, таких, что:

$$\sum_{s=1}^S \sum_{q_i, q_j \in Q_s} d_{ij} \Rightarrow \min; \quad (17)$$

$$\bigcup_s Q_s = Q, \quad Q_s \cap Q_t = \emptyset, \quad s, t = 1, 2, \dots, S; \quad (18)$$

$$|Q_s| = m_s, \quad \sum_{s=1}^S m_s = n, \quad n = |Q|. \quad (19)$$

Критерий (17) суммирует расстояния d_{ij} между объектами во всех парах $(q_i, q_j) \in Q_s$. Условия (18), (19) задают на множестве Q множество W допустимых вариантов разбиения $\{Q_s\}$. Совокупность множеств $\{Q_s\}_w$, удовлетворяющих условиям (18), (19), соответствует некоторому w -му варианту разбиения $\{Q_s\}_w \in W$.

Разбиение $\{Q_s^*\}_w$ назовем компактным разбиением (K -разбиением) множества объектов Q , если данное разбиение удовлетворяет критерию (17). Сумму расстояний d_{ij} между объектами q_i и q_j в множестве Q_s обозначим величиной L_s , т.е. $L_s = \sum_{q_i, q_j \in Q_s} d_{ij}$, а суммарную оценку компактности

для разбиения $\{Q_s\}_w$ обозначим величиной $L_w = \sum_{s=1}^S L_s$. Тогда K -разбиению

$\{Q_s^*\}_w$ будет соответствовать оценка компактности $L_w^* = \min_{w \in W} L_w$.

Применительно к задаче размещения центров, множество Q_s рассматривается как совокупность объектов, подключаемых к центру c_s , который размещен в некоторой точке q_s с координатами (x_s, y_s) . Будем считать, что координаты точки q_s для множества Q_s определяются по выражению (16). Для множества Q_s введем другую оценку компактности R_s , которая определяется относительно точки q_s как сумма расстояний d_{si} от объектов

$q_i \in Q_s$ до точки q_s , т.е. $R_s = \sum_{q_i \in Q_s} d_i$.

Аналогично оценкам L_w и L_w^* , вводится оценка $R_w = \sum_{s=1}^S R_s$ для разбиения $\{Q_s\}_w$ и оценка $R_w^* = \min_{w \in W} R_w$ для K -разбиения $\{Q_s^*\}_w$. При использовании оценок R_w критерий решения задачи разбиения запишется в виде

$$\sum_{s=1}^S \sum_{q_i \in Q_s} d_{si} \Rightarrow \min. \quad (20)$$

Относительно использования оценок L и R следует заметить, что они не противоречат друг другу, т.е. минимизация оценки L соответствует минимизации оценки R и наоборот. Также соответствие подтверждается экспериментально. В последующем будем рассматривать задачу разбиения в постановке (20), (18), (19) с вычислением координат центральных точек q_s по выражению (16).

Рассмотрим одно из возможных разбиений $\{Q_s\}_w$. В множествах Q_s , согласно (16), определены координаты точек q_s для размещения центров c_s . Решим задачу (13) – (15) подключения объектов $q_i \in Q$ к центрам в точках q_s так, чтобы на каждый центр c_s было распределено m_s объектов, а сумма расстояний, связывающих центры с распределенными на них объектами, была бы минимальной. Задачу (13) – (15) в последующем будем именовать задачей TN, обозначая тем самым то обстоятельство, что она наполовину является транспортной (Т) и наполовину задачей назначения (N).

Решение задачи TN относительно разбиения $\{Q_s\}_w$ приводит к одному из двух принципиально отличающихся результатов. В одном из них распределение объектов по центрам полностью соответствует разбиению $\{Q_s\}_w$. В этом случае на каждый центр c_s распределились объекты множества Q_s . Такое решение соответствует ситуации, когда в разбиении $\{Q_s\}_w$ объекты множества Q относительно центров c_s , $s=1, 2, \dots, S$ распределились наилучшим образом, и при этом разбиение $\{Q_s\}_w$ оказалось наилучшим относительно принятого расположения центров c_s на топологическом поле. Такое расположение центров будем именовать устойчивым. Устойчивое расположение совокупности центров $\{c_s\}$ всегда связывается с совокупностью значений $\{m_s\}$.

Другой результат решения задачи TN соответствует ситуации, когда происходит изменение некоторых или всех множеств Q_s , т.е. имеет ме-

сто перераспределение объектов между множествами. Новые множества, обозначим их Q'_s , очевидно, образуют разбиение с более высокой оценкой

компактности, т.е. $R'_{w'} \leq R_w$. Действительно, оценка $R'_{w'}$, полученная в

результате решения задачи ТН, не может быть больше оценки R_w , так как это означало бы, что решение задачи ТН не является оптимальным, т.е. хуже исходного разбиения $\{Q_s\}_w$. Следовательно, между оценками

компактности $\{Q'_s\}_{w'}$ и $\{Q_s\}_w$ должно выполняться одно из соотношений,

$R'_{w'} < R_w$ или $R'_{w'} = R_w$. Соотношение $R'_{w'} = R_w$ соответствует рассмотренной выше ситуации, при которой перераспределение объектов между множествами не произошло либо имеют место два разных разбиения, $\{Q'_s\}_{w'}$ и $\{Q_s\}_w$ с равными оценками компактности. Соотношение

$R'_{w'} < R_w$ однозначно указывает на то, что объекты между множествами Q_s перераспределились, и это привело к улучшению оценки компактности нового разбиения.

В множествах Q'_s нового разбиения остались точки q_s , установленные для множеств Q_s исходного разбиения. Эти точки не соответствуют реальным центральным точкам множеств Q'_s . Поэтому появляется возможность вычислить новые точки расположения центров c_s и улучшить

оценку компактности $R'_{w'}$. Учитывая перенос точек на новые места, разбиение $\{Q'_s\}_{w'}$ рассматривается как исходное, а задача ТН решается от-

носительно нового расположения центров c_s в надежде получить разбиение $\{Q''_s\}_{w''}$ с оценкой компактности $R''_{w''} < R_w$.

Перенос центров c_s и решение задачи ТН следует повторять до тех пор, пока вновь полученное разбиение не совпадет с предыдущим, принятым в качестве исходного. Это будет означать, что получено устойчивое расположение, которое соответствует устойчивому разбиению, названному локальным компактным разбиением (ЛК-разбиением). Найденное устойчивое разбиение названо локальным, так как нет оснований полагать, что среди множества разбиений W оно имеет наилучшую оценку компактно-

сти R_w^* , т.е. является К-разбиением.

Алгоритм решения задачи (20), (18), (19) сводится к получению ЛК-разбиения, и для этой цели используется изложенный выше метод последовательного улучшения разбиений. Для применения метода необходимо сформировать исходное разбиение и определить центральные точки в его

множествах. Исходное разбиение получается путем решения задачи ТН относительно принятой совокупности полюсов. При этом в условии (14) значение m_s уменьшается на 1, так как полюса входят в соответствующие множества Q_s . Алгоритм получения ЛК-разбиения включает следующие операции.

Задается топологическое поле с множеством объектов Q . Координаты объектов формируются автоматически по месту фиксации объектов на поле. Нумерация объектов производится в соответствии с последовательностью их на поле. Вычисляются расстояния Q_{ij} между объектами q_i и q_j , и формируется матрица расстояний D .

В множестве Q по алгоритму П1 и П2 выделяется совокупность полюсов, и относительно них на основе матрицы D формируется матрица расстояний для решения задачи ТН. Относительно полюсов решается задача ТН и определяются множества Q_s исходного разбиения.

Для каждого множества Q_s по выражению (12) определяются координаты (x_s, y_s) центральных точек q_s , и относительно них формируется новая матрица расстояний для решения задачи ТН. Решается задача ТН, и определяется новое разбиение.

Анализируется результат решения задачи ТН в п.3. Если новое разбиение отличается от исходного, то новое разбиение принимается в качестве исходного и выполняется п.3. Если новое и исходное разбиения совпадают, то это означает, что получено ЛК-разбиение.

Для решения задачи ТН был использован венгерский алгоритм. Работу алгоритма оценим на примере разбиения 35 объектов на 7 подмножеств Q_s . Исходное разбиение, полученное в результате решения задачи ТН относительно полюсов, установленных произвольным образом, имеет сумму расстояний между полюсами и подключенными к ним объектами (оценка R) – 4620 единиц. Результат применения первой итерации метода последовательного улучшения разбиений позволяет уменьшить сумму расстояний до 3461 единиц. Последующее применение метода последовательного улучшения качества разбиений приводит к результату (на 9-м шаге) с оценкой компактности ЛК-разбиения $R=2527$ единиц. Экспериментальные исследования показали, что алгоритм получает приемлемые ЛК-разбиения.

Возвращаясь к решению задачи построения локальных узлов иерархической структуры для случаев, когда на топологическом поле расположены объекты разного типа, отметим, что изложенные выше методы применимы и в данных условиях при некоторой доработке. Отличия в алгоритмах П1 и П2 связаны с тем, что после формирования очередного полюса выполняется привязка его к центру с вектором подключения, который в большей степени согласуется с типами ближайших объектов.

Что касается задачи ТН и алгоритма получения ЛК-разбиения в целом, то здесь следует отметить, что задачу ТН можно решать отдельно для объектов определенного типа. Такая возможность обусловлена тем, что объекты i -го типа могут претендовать на подключение к центрам только согласно значениям a_i в их векторах подключения. Поэтому решения, полученные для разных типов объектов при фиксированном расположении центров, оказываются независимыми друг от друга. Соответственно, оптимальное решение задачи ТН для всех типов объектов на каждой итерации алгоритма получения ЛК-разбиения складывается из оптимальных решений, полученных автономно для разных типов объектов. Это обстоятельство существенно упрощает общую задачу, так как позволяет решать ее фактически по частям с меньшей размерностью, определяемой объектами одного типа.

Таким образом, для известного числа центров на основе алгоритмов П1 и П2 и алгоритма получения ЛК-разбиения, использующего метод последовательного улучшения качества разбиений, мы получим совокупность локальных узлов иерархической структуры нижнего уровня. При необходимости, пользуясь этими же методами, можно получить совокупность локальных узлов на следующем, более высоком уровне иерархической структуры. При этом совокупность центров рассматривается в качестве нового множества терминальных объектов.

Иерархическая и многосвязная природа рассматриваемых в работе ТРС обусловила необходимость применения методов математического моделирования для решения задачи построения и оптимизации архитектуры проектируемых систем.

Представление иерархической структуры системы в виде графовой модели позволяет выполнить анализ существующих и проектирование разрабатываемых систем, исходя из условий баланса выделяемых и потребляемых коммуникационных ресурсов. На этой основе становится возможным определение оптимального числа центров, необходимых для подключения терминальных объектов, что соответствует определению рационального числа локальных узлов иерархической структуры.

Методы построения локальных узлов, сформулированные как задачи дискретного программирования, разработаны на основе выделения полюсов для размещения центров и получения локальных компактных разбиений. Эти методы и реализующие их алгоритмы эффективно работают с размерностями реальных задач и учитывают плотность расположения объектов на топологическом поле, а также наличие объектов разного типа.

Апробация предложенных методов оптимизации НТК НП проводилась для реальных распределенных систем в интересах Рослесхоза при

построении региональных многоуровневых систем мониторинга лесопожарной обстановки с использованием мобильных оперативных групп на основе БПЛА в Ханты-Мансийском автономном округе, Московской, Ростовской и Томской областях.