Заключение

1. Предложен метод анализа и классификации ЭКГ, заключающийся в вейвлет-анализе сигналов и нейросетевом распознавании образов на основе многослойного персептрона.
2. Осуществлен выбор оптимальных вариантов базисной вейвлет-функции и алгоритма обучения персептрона. Ими оказались вейвлет-функция Добеши четвертого порядка и алгоритм Левенбергам–Марквардта для обучения нейронной сети.
3. Работоспособность разработанного метода диагностирования была доказана на основе численных экспериментов.
4. Следующим этапом разработки системы будет ее оптимизация и тестирование на более длительных сигналах с целью диагностирования более широкого спектра заболеваний.

СПИСОК ЛИТЕРАТУРЫ

1. Бойцов С.А., Гришаев С.Л., Солнцев В.Н., Кудрявцев Ю.С. Анализ сигнал-усредненной ЭКГ (по данным вейвлет-преобразования) у здоровых и больных ИБС // Вестник аритмологии. – 2001. – № 23. – С. 32–34.
2. Misiti M., Misiti Y., Oppenheim G., Poggi J. Wavelets and their applications. – London: ISTE. – 2007. – 352 c.
3. Блаттер К. Вейвлет-анализ. Основы теории. – М.: Техносфера. – 2004. – 280 с.
4. Открытая база данных ЭКГ MIT. 2012. URL: http://www.physionet.org/cgi-bin/atm/ATM (дата обращения 30.01.2012).
5. Добеши И. Десять лекций по вейвлетам. – Ижевск: Регулярная и хаотическая динамика. – 2001. – 464 с.
6. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2006. – 1072 с.

Поступила 17.09.2012 г.

УДК 519.174.1

ЗАДАЧА РАЗБИЕНИЯ МНОЖЕСТВА ОБЪЕКТОВ ТЕРРИТОРИАЛЬНО РАСПРЕДЕЛЕННОЙ СИСТЕМЫ НА ПОДМНОЖЕСТВА НЕРАВНОЙ МОЩНОСТИ

Ал.В. Погребной, В.К. Погребной

Томский политехнический университет
E-mail: pogrebnoy@tpu.ru

Рассмотрена задача компактного разбиения множества объектов территориально распределенной системы на подмножества, число и мощность которых может меняться в заданных интервалах значений. Введено понятие естественных и относительных скоплений объектов. Разработан метод выделения скоплений, использующий компактные множества объектов и полученные на их основе функцию плотности и граф парных пересечений компактных множеств. Решение задачи разбиения сведено к установлению максимального соответствия между скоплениями и подмножествами.

Ключевые слова:
Задача разбиения, компактное множество, скопление объектов, функция плотности, граф парных пересечений.
Key words:
Decomposition task, compact set, collection of objects, density function, graph of the pair intersections.

Объекты территориально распределенных технических систем во многих случаях объединяются в иерархическую структуру, реализующую функции сбора данных, контроля и управления. Множество объектов на нижнем уровне иерархии разбиваются на подмножества, подключаемые к своему центру. В свою очередь, центры также могут группироваться в подмножества и подключаться к центрам более высокого уровня. Ниже исследуется задача формирования совокупности множеств объектов нижнего уровня иерархии. Принимается, что координаты размещения объектов на территории (топологическом поле [1]) известны, а фактор расстояний между объектами и центрами их подключения играет важную роль.

Задача компактного разбиения множества объектов, заданных координатами на топологическом поле, на подмножества равной мощности подробно рассмотрена в [1, 2]. Для алгоритма решения данной задачи, предложенного в [1], условие равной мощности подмножеств является существенным. Вместе с тем, во многих практических приложениях условие равной мощности подмножеств рассматривается как недостаток данного алгоритма. Очевидно, что этот недостаток в большей мере проявляется в тех случаях, когда объекты на топологическом поле расположены неравномерно. Кроме того, допускается, что число объектов подключаемых к центру, может меняться в некотором заранее установленном интервале. В этих условиях для решения задачи важно уметь выявлять компактные скопления объектов, мощность которых не выходит за пределы интервала значений мощностей, установленных для подмножеств.

Статья посвящена разработке метода решения задачи разбиения множества объектов с учетом приведенных выше условий. В основу метода положено понятие компактного множества, введенного в [2]. Метод включает выделение на топологическом поле естественных и относительных компактных скоплений объектов и формирование подмножеств в заданном интервале мощностей с минимизацией суммарной оценки компактности.

Компактные множества и функция плотности расположения объектов

Объекты множества $Q=(q_i)$, $i=1,2,...,n$, подлежащего разбиению, на топологическом поле представлены в виде точек с координатами (x_i,y_i). Согласно положениям, изложенным в [2], для объекта $q_i\in Q$ может быть сформировано множество $Q_i(g)$ с заданной мощностью g, такое, что $Q_i(g)\subset Q$, $q_i\in Q_i(g)$, а величина R_i, определяющая суммарное расстояние от центра этого множества до всех вошедших в него объектов, является минимальной. Множество $Q_i(g)$, полученное относительно объекта q_i с минимальным значением R_i, названо компактным. Алгоритм формирования компактного множества $Q_i(g)$ мощностью g заключается в выборе из множества $Q\backslash q_i$ ровно $g-1$ объектов, доставляющих минимальное значение R_i. Эффективный алгоритм формирования таких множеств предложен в [2].

Таким образом, будем считать, что для всех объектов $q_i\in Q$ получены компактные множества Q_i мощностью g. Вопросы назначения величины g и ее влияния на решение задачи разбиения требуют отдельного исследования. Предполагается, что значение g может приниматься в диапазоне от трех до величины нижней границы интервала мощностей, установленного для подмножеств разбиения. Если число центров (число подмножеств разбиения) известно и равно C, то величина g может приниматься также исходя из условия: $3\le g\le n/2C$, в котором $n/2C\ge3$. Если $n/2C<3$, то g принимается равным 3. Заметим, что независимо от принятого значения g формируется n множеств $Q_i(g)$.

Рассмотрим, каким образом можно воспользоваться наличием множеств $Q_i(g)$ для решения задачи разбиения. Напомним, что множества $Q_i(g)$ являются компактными и формируются независимо друг от друга. Следовательно множества $Q_i(g)$ могут пересекаться между собой. Увеличение мощности пересечения двух множеств $Q_i(g)$ при прочих равных условиях соответствует росту компактности (плотности) скопления объектов, входящих в эти множества. Сказанное иллюстрируется примером, представленным на рис. 1.

Суммарные оценки компактности множеств Q_1, Q_2 и Q_1', Q_2' совпадают, а мощность пересечения $|Q_1\cap Q_2|>|Q_1'\cap Q_2'|$. Это соответствует тому, что компактность объединения $(Q_1\cup Q_2)$ выше (значение оценки R меньше), чем у объединения $(Q_1'\cup Q_2')$. Данные объединения представляют соответствующие локальные скопления, компактность которых можно сопоставлять с плотностью расположения объектов на топологическом поле.

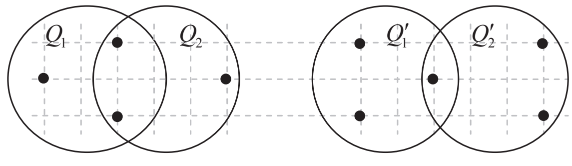

Рис. 1. *Сравнение мощности пересечения множеств и компактности скоплений*

В общем случае будем сравнивать две совокупности множеств $Q_i(g)$, обозначим их A и B, каждая из которых содержит равное число $Q_i(g)$ сформированных для определенного значения g. Пусть в A суммарная мощность парных пересечений множеств $Q_i(g)$ составляет величину $\rho(A)$, а в B – $\rho(B)$. Мощность объединения объектов множеств $Q_i(g)$ в A обозначим $s(A)$, а в B – $s(B)$. Если $\rho(A)>\rho(B)$, то нетрудно установить, что $s(A)<s(B)$. С увеличением $\rho(A)$ уменьшается число объектов, образующих некоторое скопление, на котором сформированы множества совокупности A, что соответствует увеличению плотности этого скопления. Из этого следует, что совокупность A объединяет множество объектов, которое относительно своего окружения выделяется как более плотное скопление.

Наличие связи параметра пересечений множеств $Q_i(g)$ с относительной плотностью скоплений объектов имеет принципиально важное значение для выявления скоплений и формирования на этой основе подмножеств разбиения. Участие объекта $q_i\in Q$ в парных пересечениях множеств $Q_i(g)$, сформированных для фиксированного g, представим функцией $F(g)$. Значение $F_i(g)$ функции $F(g)$ равно числу вхождений объекта q_i в множества $Q_i(g)$. Функцию $F(g)$ назовем функцией плотности скоплений объектов множества Q на топологическом поле. Принадлежность объекта q_i к пересечению некоторой совокупности компактных множеств свидетельствует о том, что в окрестности q_i присутствуют объекты, которые наряду с q_i входят в эту совокупность компактных множеств. Например, значение $F_i(g)=5$ означает, что объект q_i входит в состав пяти компактных множеств. Учитывая, что каждое из пяти множеств является компактным, можно говорить о принадлежности q_i к некоторому скоплению объектов, образованному этими множествами.

Выделение естественных компактных скоплений объектов

Неравномерная плотность расположения объектов на топологическом поле предполагает наличие скоплений, в которых объекты расположены более плотно. Пространство топологического поля между скоплениями может быть пустым либо содержать объекты удаленные друг от друга на расстояния, которые соответствуют более низкой плотности, чем в соседних скоплениях. Здесь важно, что скопления выделяются как совокупность объектов удаленных от объектов окружения на расстояния, превышающие расстояния между объектами внутри скопления. Скопления, которые располагаются на топологическом поле и выделяются независимо от решения задачи разбиения множества объектов на подмножества, назовем *естественными*. Совокупность объектов S из множества Q является естественным скоплением, если все компактные множества $Q_i(g)$ мощностью g, сформированные для $q_i \in S$, состоят из объектов совокупности S, т. е. $Q_i(g) \subseteq S$.

Выделение естественных скоплений (ЕС) рассмотрим на примере фрагмента топологического поля, представленного на рис. 2.

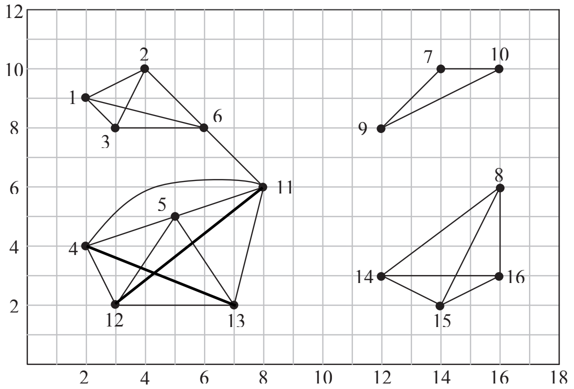

Рис. 2. *Фрагмент топологического поля*

Множество Q в примере на рис. 2 содержит 16 объектов. Ярко выраженные ЕС на топологическом поле отсутствуют. Визуально можно выделить три ЕС: S_1=(1, 2, 3, 4, 5, 6, 11, 12, 13), S_2=(7, 9, 10), S_3=(8, 14, 15, 16). Рассмотрим возможные подходы к формализации методов решения задачи выделения ЕС. Примем g=3 и сформируем компактные множества $Q_i(3)$. Состав множеств $Q_i(3)$ и значения функции $F_i(3)$ представим в виде таблицы.

q_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$Q_i(3)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	2	1	1	5	4	2	9	15	7	7	5	4	5	15	14	14
	3	3	2	12	12	3	10	16	10	9	6	5	12	16	16	15
$F_i(3)$	3	4	4	3	5	2	3	1	3	3	1	4	1	3	4	4

Каждое множество $Q_i(3)$ представлено в столбце таблицы тремя объектами. На первом месте вверху столбца указан объект q_i, относительно которого сформировано компактное множество $Q_i(3)$. Внизу таблицы приведены значения функции $F_i(3)$, полученные для соответствующих q_i. Например, $F_6(3)$=2, т. к. объект q_6 входит в $Q_6(3)$ и $Q_{11}(3)$. Заметим, что компактные множества $Q_i(3)$ сформированы по алгоритму [2] согласно координатной сетке, обозначенной на рис. 2 снизу и слева от топологического поля. Оценки компактности R_i множеств $Q_i(3)$ вычислялись при этом относительно центров, которые определялись как центры единичных масс.

Используя значения $F_i(g)$, после выполнения аппроксимации получим функцию $F_i^*(g)$ в виде «холмистой» поверхности, на которой каждый «холм» соответствует скоплению объектов. Если все значения $F_i(g)$ уменьшить на величину $\min_i\{F_i(g)\}$, то некоторые «холмы» на поверхности $F^*(g)$ предстанут в виде «островов». По границе «острова» устанавливается граница соответствующего скопления. Выделенные скопления исключаются из множества Q, а операция вычитания минимального значения $F_i(g)$ и исключения скопления при необходимости повторяется. Объекты, которые оказались «потерянными» при выполнении операции вычитания минимального значения $F_i(g)$ распределяются между скоплениями.

Для рассматриваемого примера после вычитания 1 из $F_i(g)$ теряются соответствующие объекты q_8, q_{11}, q_{13}, что приводит к обособлению четырех «островков» (скоплений) S_1^*=(1,2,3,6), S_2^*=(4,5,12), S_3^*=(7,9,10), S_4^*=(14,15,16). Заметим, что компактные множества, сформированные на объектах в разных скоплениях, не пересекаются между собой. По этому признаку, в частности, можно определять границы скоплений. Распределение «потерянных» объектов q_8, q_{11}, q_{13} по скоплениям производится в соответствии с их минимальным удалением от объектов скоплений: q_8 в S_4^*, q_{13} в S_2^*, q_{11} в S_3^*.

Полученную совокупность скоплений можно рассматривать как результат решения задачи свободного разбиения множества Q на подмножества, каждое из которых соответствует определенному скоплению. Очевидно, что свободное разбиение не всегда соответствует требованиям к разбиению по количеству подмножеств и их мощности, поэтому выделение ЕС и скоплений с уточненными границами будем рассматривать как этап предварительного анализа характера расположения объектов множества Q на топологическом поле.

Формирование подмножеств в заданном интервале мощностей

Основные параметры разбиения, связанные с ограничением числа подмножеств C и наличием интервала изменения допустимых мощностей $[a,b]$, могут устанавливаться в зависимости от результатов выделения скоплений. Решение задачи разбиения для заданных параметров будем рассматривать независимо от того получены они на основе выделенных скоплений или назначены по другим основаниям. Вместе с тем при разработке метода разбиения будем стремиться максимально использовать полученные скопления в качестве подмножеств.

Введем граф парных пересечений компактных множеств и обозначим его G. Вершины графа G соответствуют компактным множествам Q_i, а ребра (Q_i,Q_j) – наличию пересечений $Q_i \cap Q_j \neq \varnothing$ с весом $r_{ij}=|Q_i \cap Q_j|$. Граф G, отражающий парные пересечения компактных множеств для рассматриваемого примера, приведен на рис. 2. Вершины графа соответствуют объектам q_i, относительно которых были построены соответствующие компактные множества Q_i.

Матрица связности вершин данного графа с указанием весов ребер r_{ij} представлена на рис. 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		3	3			2											1
2	3		3			2											2
3	3	3				2											3
4					3						1	3	2				4
5				3							1	3	2				5
6	2	2	2								1						6
7									3	3							7
8														2	2	2	8
9							3			3							9
10							3		3								10
11				1	1	1						1	1				11
12				3	3						1		2				12
13				2	2						1	2					13
14								2							3	3	14
15								2						3		3	15
16								2						3	3		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Рис. 3. *Пример матрицы связности вершин графа G*

Граф на рис. 2 состоит из трех компонент связности [3], которые по составу объектов соответствуют скоплениям S_1, S_2, S_3, выделенным ранее визуально. В условиях, когда граф G является связным, т. е. состоит из одной компоненты связности, выделение скоплений можно осуществлять путем удаления ребер с минимальным весом, по аналогии с тем, как это делалось ранее при вычитании минимальных значений функций плотности. Например, удаление пяти ребер с весом r_{ij}=1, связывающих вершину Q_{11} в скоплении s_1=(1,2,3,4,5,6,11,12,13), приводит к выделению скоплений (1,2,3,6) и (4,5,12,13). Полученная совокупность из четырех скоплений в данном примере может служить основой для принятия решения о разбиении множества Q на C=4 подмножества с интервалом мощностей $[a,b]$=[3,5]. Для данных значений параметров C и $[a,b]$ после включения объекта q_{11} в скопление (4,5,12,13) имеем решение задачи разбиения. В общем случае, когда требуемые значения параметров C и $[a,b]$ назначены без предварительного выделения и анализа скоплений, решение задачи разбиения множества Q на подмножества включает выполнение следующих этапов.

1. По заданным значениям C и $[a, b]$ определяются усредненные (исходные) мощности подмножеств разбиения. Для этого величина n представляется в виде суммы (разложения) $n_0(C-k)+(n_0+1)k$. Здесь n_0 – целая часть величины n/C, а $k=n-n_0C$. Разложение величины n предполагает разбиение Q на $(C-k)$ подмножеств мощностью n_0 и k подмножеств мощностью (n_0+1). Например, при n=32 и C=5 получим n_0=6, k=2, и мощности подмножеств будут приниматься согласно разложению (6,6,6,7,7). Для примера на рис. 2 при C=3 разложение n=16 получим в виде (5,5,6), а интервал $[a,b]$ в этом случае может быть принят равным [4,7].
2. Для величины g, взятой из интервала $[3, a]$, формируются компактные множества $Q_i(g)$. При выборе исходного значения g предпочтение отдается $g=a$. Для проведения более глубокого анализа наличия ЕС меньшей мощности значение g может последовательно понижаться до 3.
3. На основе анализа функции плотности $F(g)$ или (и) выделения компонент связности графа $G(g)$ проверяется наличие ЕС. Все ЕС, мощность которых укладывается в интервал $[a,b]$, заносятся в разложение, замещая его элементы. После этого незамещенные элементы разложения корректируются так, чтобы сумма элементов была равна n. Если ЕС отсутствуют, т. е. граф $G(q)$ состоит из одной компоненты связности или мощность ЕС превышает b, то соответствующая компонента связности подлежит разбиению.
4. Разбиение компонент связности и графа G выполняется с помощью матричного алгоритма разрезания графов на минимально связанные подграфы и соответствующего программного средства CutGraf [4]. При этом размерность блоков матрицы-шаблона принимается в соответствии с элементами исходного разложения, если разрезается граф G. Если разрезается компонента связности, то используются незамещенные элементы из скорректированного разложения.
5. Для выделенных ЕС и скоплений, полученных в результате решения задачи разрезания, определяются центры, и относительно них решается задача подключения объектов к центрам по критерию минимизации суммарной оценки компактности скоплений (подмножеств). Задача формулируется как задача математического программирования транспортного типа [5]. Необходимость решения данной задачи возникает при наличии «потерянных» объектов либо при желании оценить качество решения задачи разбиения.

Решение задачи разбиения для примера на рис. 2 с заданными параметрами C=4, $[a,b]$=[3,5] приводит к разложению (4,4,4,4). После формирования $Q_i(3)$, построения и анализа графа $G(3)$, выделяются три ЕС (три компоненты связности). Две из них (7,9,10) и (8,14,15,16) укладываются в интервал и заносятся в разложение, замещая два последних элемента. В результате корректировки получается новое разложение (4,5,3,4). На рис. 4 приведены результаты разрезания графа $G(3)$ для размерностей блоков матрицы-шаблона, соответствующих элементам скорректированного разложения. Справа и снизу матрицы приведены номера объектов, вошедших в блоки, которые соответствуют подмножествам разбиения.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		3	3	2													1
2	3		3	2													2
3	3	3		2													3
4	2	2	2			1											6
5						1	3	2	3								5
6				1	1		1	1	1								11
7					3	1		2	3								4
8					2	1	2		2								13
9					3	1	3	2									**12**
10											3	3					10
11										3		3					7
12										3	3						**9**
13														2	2	2	8
14													2		3	3	14
15													2	3		3	15
16													2	3	3		16
	1	2	3	6	5	11	4	13	**12**	10	7	**9**	8	14	15	16	

Рис. 4. *Результат разрезания графа G(3)*

При назначении параметров C=3 и $[a,b]$=[4,7] получим разложение (5,5,6). Граф G (4), построенный для $Q_i(4)$, состоит из двух ЕС. Одно из них (7,8,9,10,14,15,16) укладывается в интервал, второе разрезается согласно скорректированному разложению (4,5,7). Оба варианта решения задачи разбиения не приводят к появлению «потерянных объектов».

Программная реализация перечисленных этапов решения задачи разбиения предполагает более детальную алгоритмизацию этапов и в ряде случаев интерактивное вмешательство пользователя. Необходимость вмешательства возникает при выборе величины g, выделении ЕС и скоплений на основе функции $F(g)$ или графа $G(g)$, применении алгоритмов решения задачи подключения объектов к центрам и в ряде других случаях.

Выводы

1. Замена объектов на компактные множества и соответственно отношений между объектами в виде расстояний на отношения между компактными множествами в виде мощностей парных пересечений усиливает эффект локализации скоплений объектов на топологическом поле.
2. Использование эффекта локализации скоплений и введение функции плотности и графа парных пересечений позволило разработать эффективный инструмент выделения скоплений. На топологическом поле скопления могут присутствовать изначально в автономном виде (естественные скопления) или как скопления объектов, которые относительно своего ближайшего окружения выделяются большей плотностью (относительные скопления).
3. При наличии инструмента выделения скоплений задача разбиения множества объектов на подмножества в заданном интервале мощностей в основном сводиться к нахождению максимального соответствия между скоплениями и подмножествами с учетом их числа и допустимой мощности.
4. Понятие скопления, введенное в данной статье, остается малоизученным. Исследование свойств скоплений и введение соответствующих оценок позволит в последующем разработать алгоритмы решения задач разбиения с учетом конкретных требований для других практических приложений.

Работа выполнена в рамках госзадания «Наука»

СПИСОК ЛИТЕРАТУРЫ

1. Погребной Ал.В., Погребной Ан.В. Алгоритм решения задачи компактного разбиения множества объектов территориально распределенной системы // Известия Томского политехнического университета. – 2010. – Т. 317. – № 5. – С. 22–28.
2. Погребной Ал.В. Определение нижней границы оценки компактности для топологических разбиений // Известия Томского политехнического университета. – 2011. – Т. 319. – № 5. – С. 27–32.
3. Кристофидес Н. Теория графов. Алгоритмический подход. – М.: Мир, 1978. – 432 с.
4. Погребной Ан.В., Погребной Д.В. Исследование матричного алгоритма решения задачи разрезания графов // Молодежь и современные информационные технологии: Труды VIII Всеросс. научно-практ. конф. молодых ученых. – Томск, 2010. – С. 140–148.
5. Гольштейн Е.Г., Юдин Д.Б. Задачи линейного программирования транспортного типа. – М.: Наука, Физматлит, 1969. – 384 с.

Поступила 27.09.2012 г.