

## Агрегирование системы правил в продукционных экспертных системах

НТУ «ХПИ»

**Ключевые слова:** экспертная система, продукционные правила, агрегирование, принцип максимума правдоподобия

**Ключові слова:** експертна система, продукційні правила, агрегування, принцип максимуму правдоподібності

**Keywords:** expert system, products instiled, aggregation, maximum likelihood principle

Введение Традиционные технологии построения экспертных систем диагностики состояния объектов основаны на использовании результатов контроля наблюдаемого набора контролируемых параметров объекта. При этом наиболее простая и естественная процедура обработки совокупности наблюдений реализуется в так называемых продукционных экспертных системах. Как известно [1,2], продукционные экспертные системы диагностирования состояний объекта используют систему правил, каждое из которых состоит из двух частей: посылка (условие-антецедент) и заключение (следствие-консеквент). Если при этом контролируется  $m$  параметров объекта, причем  $i$ -й параметр может принять одно из  $l_i$  возможных значений, то общее число  $R$  продукционных правил, необходимое для диагностики всех состояний объекта, определяется соотношением  $R = l_1 l_2 \dots l_m$ . Таким образом, эта система правил образует  $R$ -компонентный вектор, размерность которого определяется числом контролируемых параметров.

Экспоненциальный характер зависимости числа продукционных правил от количества контролируемых параметров приводит к необходимости модернизации стандартной технологии построения продукционных систем.

Цель статьи – разработка технологии радикального снижения размерности задачи диагностики состояния объектов с большим числом контролируемых параметров.

Постановка задачи Реальное снижение размерности задачи диагностики возможно осуществить по следующим двум направлениям. Во-первых, при решении многих практических задач диагностики состояния объектов может быть достаточно использовать информацию только о наличии или отсутствии некоторых признаков (симптомов), являющихся атрибутами состояний. Во-вторых, для решения задачи диагностики во многих случаях приемлемые результаты могут быть получены при использовании агрегированной по тем или иным правилам информации о параметрах объекта.

Поставим задачу оценки эффективности перечисленных направлений преодоления «проклятия» размерности.

### Основные результаты

1. Рассмотрим возможные варианты построения ЭС, реализующих перечисленные принципы снижения размерности задачи.

Первый вариант. Введем индикатор

$$\eta_i = \begin{cases} 1, & \text{если } i\text{-й симптом наблюдается в результате контроля,} \\ 0 & \text{– в противном случае.} \end{cases}$$

Тогда каждый результат комплексного контроля  $m$  симптомов пациента представляет собой набор  $\eta = \{\eta_1, \eta_2, \dots, \eta_m\}$ .

При этом необходимое для исчерпывающего описания всех возможных ситуаций число правил  $R$  будет равно  $2^m \ll l_1 l_2 \dots l_m$ . Однако ясно, что существенного снижения общего числа правил не происходит.

Второй вариант приводит к радикальному упрощению процедуры диагностики. Введем

$I = \{1, 2, \dots, m\}$  – множество контролируемых симптомов,

$k = \{1, 2, \dots, K\}$  – множество возможных состояний объекта диагностики,

$l_{ik} = \begin{cases} 1, & \text{если } i\text{-й симптом типичен для } k\text{-го состояния,} \\ 0 & \text{противном случае.} \end{cases}$

Вычислим

$n_{11}^{(k)} = \sum_{i=1}^m l_{ik} \eta_i$  – число типичных для  $k$ -го состояния объекта симптомов, наблюдаемых по результатам контроля;

наблюдаемых по результатам контроля;

$n_{00}^{(k)} = \sum_{i=1}^m (1 - l_{ik})(1 - \eta_i)$  – число нетипичных для  $k$ -го состояния объекта симптомов, не наблюдаемых по результатам контроля;

симптомов, не наблюдаемых по результатам контроля;

$n_{10}^{(k)} = \sum_{i=1}^m l_{ik}(1 - \eta_i)$  – число типичных для  $k$ -го состояния объекта симптомов, не наблюдаемых по результатам контроля;

симптомов, не наблюдаемых по результатам контроля;

$n_{01}^{(k)} = \sum_{i=1}^m (1 - l_{ik})\eta_i$  – число нетипичных для  $k$ -го состояния объекта симптомов, наблюдаемых по результатам контроля.

симптомов, наблюдаемых по результатам контроля.

Определим теперь

$n_{np}^{(k)} = n_{11}^{(k)} + n_{00}^{(k)}$  – число правильных ситуаций (типичные симптомы наблюдаются, а нетипичные – нет), для  $k$ -го состояния объекта, возникших в результате контроля;

$n_{л}^{(k)} = n_{01}^{(k)} + n_{10}^{(k)}$  – число неправильных, ложных ситуаций (нетипичные симптомы наблюдаются, а типичные – нет), для  $k$ -го состояния объекта, возникших в результате контроля.

Тогда численное значение показателя

$$p^{(k)} = \frac{n_{11}^{(k)} + n_{00}^{(k)}}{n} = \frac{n_{np}^{(k)}}{n}, \quad k=1, 2, \dots, K, \quad (1)$$

характеризует степень уверенности в правильности распознавания  $k$ -го состояния объекта. Этот и некоторые другие варианты расчета критерия распознавания описаны в [3] и приведены в таблице 1

Таблица. Критерии распознавания

№ п/п	Авторы	Расчетная формула
1	Кульжинский	$p_1 = \frac{n_{11} + n_{00}}{n} = \frac{n_{np}}{n}$
2	Кульжинский	$p_2 = \frac{n_{11}n}{2(n_{11} + n_{10})(n_{11} + n_{01})}$
3	Кульжинский	$p_3 = \frac{n_{11}}{(n_{11} + n_{10})(n_{11} + n_{01})}$
4	Рассел, Рао	$p_4 = \frac{n_{11}}{n}$
5	Хамман	$p_5 = \frac{n_{np} - n_l}{n}$
6	Джакарт	$p_6 = \frac{n_{11}}{n_{11} - n_{10} - n_{01}}$
7	Дейк	$p_7 = \frac{2n_{11}}{2n_{11} - n_{10} - n_{01}}$
8	Сокаль, Сниф	$p_8 = \frac{n_{11}}{2n_{11} - n_{10} - n_{01}}$
9	Мушталир	$p_9 = \frac{n_{np}}{n - n_l}$
10	Жакар, Нидман	$p_{10} = \frac{n_{11}}{n_{11} - n_l}$
11	Юл	$p_{11} = \frac{n_{11}n_{00} - n_{10}n_{01}}{n_{11}n_{00} + n_{10}n_{01}}$

Здесь для упрощения записи индекс  $k$  опущен. Приведенные в таблице 1 критерии в неравной мере обеспечивают решение задачи диагностики состояния. В частности, критерий  $p_4$  не учитывает важной информации об отсутствии нетипичных симптомов. Критерий  $p_6$ , принимая правильное значение, равное 1 при  $n_l = 0$ , с ростом  $n_l$  в интервале  $(0, n_{11})$  неограниченно растет, а затем становится отрицательным. Аналогично ведут себя критерии  $p_7$  и  $p_8$ . Численные значения критериев  $p_1$  и  $p_5$  линейно возрастают с увеличением

числа правильных ситуаций, а для критериев  $p_2, p_3, p_9, p_{10}, p_{11}$  этот рост имеет нелинейный, что трудно объяснимо, характер. В связи с этим содержательная их трактовка не вполне ясна. На практике наиболее употребителен критерий  $p_1$ , задающий долю правильных ситуаций, возникающих при диагностике каждого из возможных состояний объекта.

Принципиальное достоинство приведенных схем диагностики состояния состоит в фактической агрегации результатов контроля симптомов, что практически снимает проблему размерности задачи. Вместе с тем, все приведенные критерии обладают общим недостатком: они не учитывают возможные существенные различия в информационной ценности симптомов. Причина этого состоит в неопределенности терминов «типичный симптом» и «нетипичный симптом».

Осуществим учет информационной ценности симптомов. С целью получения формального описания этих терминов, допускающего ясную трактовку, введем вероятность  $P_{ik}$  появления  $i$ -го симптома при условии, что объект находится в  $k$ -м состоянии. По сути  $P_{ik}$  есть вероятность того, что  $i$ -й симптом является типичным для  $k$ -го состояния объекта.

Теперь, с использованием набора  $\{P_{ik}\}, i=1,2,\dots,m, k=1,2,\dots,K$ , вычислим:

$$\tilde{n}_{11}^{(k)} = \sum_{i=1}^m p_{ik} \eta_i - \text{среднее число наблюдений симптомов, типичных для } k\text{-го состояния};$$

го состояния;

$$\tilde{n}_{00}^{(k)} = \sum_{i=1}^m (1 - p_{ik}) \eta_i - \text{среднее число наблюдений симптомов, нетипичных для } k\text{-го состояния};$$

для  $k$ -го состояния;

$$\tilde{n}_{10}^{(k)} = \sum_{i=1}^m p_{ik} (1 - \eta_i) - \text{среднее число ненаблюдений симптомов, типичных для } k\text{-го состояния};$$

для  $k$ -го состояния;

$$\tilde{n}_{01}^{(k)} = \sum_{i=1}^m (1 - p_{ik}) (1 - \eta_i) - \text{среднее число ненаблюдений симптомов, нетипичных для } k\text{-го состояния.}$$

Далее

$$\tilde{n}_{np}^{(k)} = \tilde{n}_{11}^{(k)} + \tilde{n}_{00}^{(k)} - \text{среднее число правильных ситуаций для } k\text{-го состояния объекта};$$

состояния объекта;

$$\tilde{n}_n^{(k)} = \tilde{n}_{01}^{(k)} + \tilde{n}_{10}^{(k)} - \text{среднее число неправильных ситуаций для } k\text{-го состояния объекта.}$$

состояния объекта.

Тогда теоретико-вероятностный аналог критерия распознавания (1) будет иметь вид

$$\tilde{p}_1^{(k)} = \frac{\tilde{n}_{np}^{(k)}}{n}, \quad k = 1, 2, \dots, K.$$

Аналогично этому могут быть рассчитаны и другие критерии распознавания состояния объекта, приведенные в таблице. Понятно, что степень информативности каждого из симптомов при диагностике конкретного состояния будет высока, если он часто проявляется в случаях, когда объект находится в этом состоянии, и, одновременно, проявляется редко, если объект находится в каком-либо другом состоянии. Этот же симптом будет мало информативен, если его наличие или отсутствие слабо связаны с состоянием объекта.

2. Построение продукционной ЭС с использованием принципа максимума правдоподобия.

Ввод набора вероятностей  $\{P_{ik}\}$ , позволяет в полной мере реализовать безупречную конструкцию продукционных правил, обеспечивая возможность расчета вероятностей правильности делаемых заключений. С этой целью для каждого из возможных состояний объекта по результатам контроля  $\eta = \{\eta_1, \eta_2, \dots, \eta_m\}$  рассчитаем

$$P^{(k)} = \prod_{i=1}^m P_{ik}^{\eta_i} (1 - P_{ik})^{1-\eta_i}, \quad k = 1, 2, \dots, K.$$

Значение  $P^{(k)}$ , как легко видеть, имеет смысл правдоподобия  $k$ -го состояния при условии, что наблюдаемым является набор  $\eta = \{\eta_1, \eta_2, \dots, \eta_m\}$ .

Вычисление набора значений  $P^{(k)}$ ,  $k = 1, 2, \dots, K$  завершает построение набора продукционных правил. При этом количество продукционных правил доведено до потенциального минимума – числа возможных диагнозов.

Теперь важно отметить, что использование введенной технологии расчета вероятностей правдоподобия диагноза для любого из результатов контроля позволяет радикально изменить принцип функционирования продукционных экспертных систем, так как снимает необходимость формирования системы продукционных правил, устраняя, таким образом, «проклятие» размерности.

**Выводы.** Описанная выше технология диагностирования состояния объекта может быть распространена на случай, когда в качестве исходной информации используются не булевы переменные-симптомы, а измеряемые параметры, принимающие значения в некотором диапазоне. Понятно, что точность диагностики при этом повышается. Недостаток предложенной методики оценивания состояния состоит в том, что численные значения рассчитываемых вероятностей состояний отыскиваются независимо друг от друга. Этот недостаток устраняется при использовании байесова механизма логического вывода.

### Список литературы

1. Уотермен Д. Руководство по экспертным системам: пер. с англ./ Д.Уотермен. – М.: МИР, 1989.–388 с.
2. Нейлор К. Как построить свою экспертную систему: пер. с англ./ К.Нейлор. – М.: Энергоатомиздат, 1991.–288 с.
3. Елисеева И.И. Группировка, корреляция, распознавание образов/И.И.Елисеева, В.С. Рукавишников. – М.:Статистика, 1997. – 144 с.

**Рецензент:**, д.т.н., проф. Л.Г. Раскин, НТУ «ХПИ», г. Харьков

Поступила в редакцию 13.03.09