АЛГОРИТЪМ ЗА ВЕРОЯТНОСТНИ РАЗСЪЖДЕНИЯ В ИНТЕЛИГЕНТНИТЕ ДИАГНОСТИЧНИ СИСТЕМИ

Марчела Банова, Технически университет- София

1. ВЪВЕДЕНИЕ.

В моделно-базираните системи процесът на диагностика преминава през две фази [1]. Целта на първата фаза е да се определят множествата от дефектирали компоненти (наречени диагнози), които обясняват наблюдаваните симптоми. Целта на втората фаза е да се направят допълнителни измервания, които най-добре разграничават отделните възможни диагнози.

Процесът на определяне на допълнителните измервания зависи силно от наличните вероятности за дефектиране на отделните компоненти на диагностираното устройство.

В настоящата статия се предлага алгоритъм за определяне на оптималната съвкупност от допълнителни измервания, като компонентите на диагностираното устройство се разделят на групи. Всички елементи от една група дефектират с еднаква вероятност или с много по-голяма или по-малка вероятност в сравнение с елементите от другите групи.

На базата на това алгоритъмът определя вероятностите за наличие на всяка от предполагаемите диагнози. Диагностичната стратегия [3] цели първо да се направят измервания, които позволяват диференциране на еднократните диагнози. Едва след елиминиране на всички еднократни диагнози, се определят необходимите измервания за диференциране на двукратните диагнози и т.н.

В диагностичния алгоритъм се прилага подхода на минималната ентропия [2].

2. МАТЕМАТИЧЕСКИ МОДЕЛ.

Предлаганият вероятностен алгоритъм е разработен на базата на следните допускания:

(i) Компонентите на диагностираното устройство дефектират независимо.

Това допускане е широко застъпено в моделно-базираните диагностични системи, въпреки че не винаги е вярно.

(ii) Множеството от компонентите е разделено на отделни подмножества D_1, D_2, ... , D_k. Всички елементи на множеството D_i дефектират с една и съща вероятност $\mathcal{E}_i$, която е съществено по-голяма или по-малка от вероятностите за дефектиране на компонентите от другите множества D_j.

Това допускане е оправдано в случая, когато от статистически данни е известно, че напр. транзисторите дефектират с по-голяма вероятност от резисторите и т.н.

При тези допускания априорната вероятност p_ℓ диагнозата C_ℓ да е коректна се дава с израза

(1) $p_\ell = \prod_{c\in C\ell} p(c\in C_a) \cdot \prod_{c\notin C\ell} [1-p(c\in C_a)]$,

където: $C_\ell = \{c_1, c_2, ..., c_\ell\}$ - диагноза

C_a - множество от дефектирали компоненти.

Да отбележим, че:

$C_a = D_1 \cup D_2 \cup ... \cup D_k$

$p(c\in D_i) = \mathcal{E}_i$, i=1,...,k

Нека да означим с t_i броя на елементите, принадлежащи на D_i, които са елементи на диагнозата C_ℓ. Така на всяка диагноза C_ℓ се съпоставя един вектор $(t_1, t_2, ..., t_k)$ и изразът (1) придобива вида:

(2) $p_\ell = \varepsilon_1^{t_1} \cdot \varepsilon_2^{t_2} \dots \varepsilon_k^{t_k} \cdot (1-\varepsilon_1)^{|D_1|-t_1} \dots (1-\varepsilon_k)^{|D_k|-t_k}$

Ако отделните вероятности ε_i са много малки, априорната вероятност p_ℓ може да се апроксимира с

(3) $p_\ell \approx \varepsilon_1^{t_1} \cdot \varepsilon_2^{t_2} \dots \varepsilon_k^{t_k} = \prod_{i=1}^{k} \varepsilon_i^{t_i}$

В случая, когато е направено допълнително измерване и е измерена стойност V_{ik} в контролната точка X_i, условната вероятност на диагнозата C_ℓ се получава от Правилото на Бейс:

(4) $$p(C_\ell / x_i = V_{ik}) = \frac{p(x_i = V_{ik} / C_\ell) \cdot p(C_\ell)}{p(x_i = V_{ik})}$$

където: $p(x_i = V_{ik})$ - нормализация

$$p(x_i = V_{ik} / C_\ell) = \begin{cases} 1, \text{ ако } C_\ell \Longrightarrow x_i = V_{ik} \\ 0, \text{ ако } C_\ell \not\Longrightarrow x_i = V_{ik} \\ \quad 1/2, \text{ ако } C \text{ не предсказва} \end{cases}$$

След Е на брой измервания вероятността диагнозата C_ℓ да не бъде елеминирана е:

(5) $p(C_\ell / E) = \prod_{i=1}^{k} \varepsilon_i^{t_i} / N \cdot 2^{f_\ell}$

където: f_ℓ - брой пъти, когато диагнозата C_ℓ предсказва измерената в контролната точка стойност

N - нормализация

Очакваната апостериорна ентропия нна диагнозите след като е направено измерване в контролната точка X_i се дава [2] с израза:

(6) $\sum_{k=1}^{m_i} [p(S_{ik})+p(U_i)/2] \cdot \ln[p(S_{ik})+p(U_i)/2] - p(U_i) \cdot \ln 1/2$

където: S_{ik} - множеството от диагнози, които предсказват, че в к.т. X е измерена стойност V_{ik}

U_i - множеството от диагнози, които не предсказват измерената в к.т. X_i стойност.

За да се определи най-доброто измерване (което ще елиминира най-голям брой диагнози) трябва да се минимизира изразът (6) върху всички к.т. X_i.

Изразът (6) може да се опрости като се използва факта, че сумата от вероятностите на всички диагнози е 1 (разглеждат се само диагнозите с мин. кратност), т.е.

$$(7)\quad 1 = \sum_{C_j} p(C_\ell/E) = \sum_{C_j} \frac{\prod \varepsilon_i^{t_i}}{N.2^{f\ell}}$$

където: C_j - диагнозите с минимална кратност

Следователно:

$$(8)\quad N = \sum_{C_j} \frac{\prod \varepsilon_i^{t_i}}{2^{f\ell}} = \prod^{k} \varepsilon_i^{t_i} \cdot \sum_{C_j} \frac{1}{2^{f\ell}}$$

От (8) и (5) следва:

$$(9)\quad p(C_\ell/E) = 1/2^{f\ell} \cdot \sum_{C_j} \frac{1}{2^{f\ell}}$$

Както се вижда, изразът (9), а следователно и изразът (6), не зависи от ε_i, т.е. (6) е проста функция на f_ℓ.

Ако допуснем, че диагнозите винаги предсказват определена стойност за всяка от контролните точки, т.е. ако:

$p(U_i) = 0$ и

$f_i = 0$,

изразът (6) се опростява до

$$(10)\quad \sum_{K=1}^{m} C_{iK}/N' \ln \left[C_{iK}/N'\right]$$

където: C_{iK} - брой на q-кратните диагнози, предсказващи, че $X_i = V_{iK}$

N' - общият брой q-кратни диагнози

Тъй като N' е константа за всички контролни точки, изразът (10) (а следователно - и (6)) се опростява до

$$(11)\quad \sum_{K=1}^{m} C_{iK} . \ln C_{iK}$$

Минимизирането на израза (11) не представлява изчислителен проблем.

3. ОПИСАНИЕ НА АЛГОРИТЪМА.

На базата на получените и представени в т.2 теоретични резултати е разработен вероятностен алгоритъм, който е съставна част на диагностичния алгоритъм, използван в интелигентната диагностична система DIAN [3], както е показано на Фиг.1.

След въвеждане на диагностичните симптоми, алгоритъмът за диагностика на базата на първия принцип CARD_DIAGNOSE определя оптималното множество от правдобододобни диагнози. След това се прилага вероятностният алгоритъм, който използва допълнителни измервания за да елиминира (и съответно - да потвърди) отделни диагнози.

4. ЗАКЛЮЧЕНИЕ

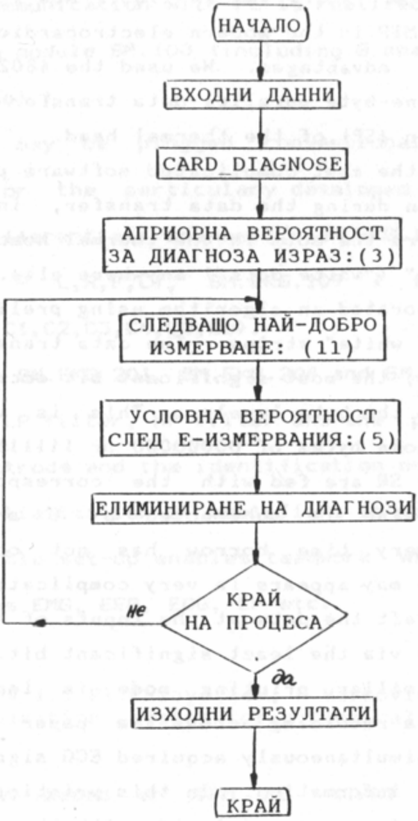
Предимство на предлагания вероятностен алгоритъм е неговата изчислителна простота. Недостатък на алгоритъма е, че не винаги измерванията, разграничаващи две еднократни диагнози, разграничават диагнозите от по-висока кратност, т.е. - правят се по-голям брой измервания. Алгоритъмът е реализиран в интелигентната диагностична система DIAN.

БИБЛИОГРАФИЯ

[1] R.Reiter, A theory of diagnosis from first principles, Artificial Intelligence 32 (1987), p.57-95

[2] J.de Kleer, B.C.Williams, Diagnosing multiple faults, Artificial Intelligence 32 (1987), p.97-130

[3] М.Банова, DIAN-интелигентна система за диагностика (подход за диагностика), 28 научна сесия КЕКС 1993


НАЧАЛО
ВХОДНИ ДАННИ
CARD_DIAGNOSE
АПРИОРНА ВЕРОЯТНОСТ ЗА ДИАГНОЗА ИЗРАЗ:(3)
СЛЕДВАЩО НАЙ-ДОБРО ИЗМЕРВАНЕ: (11)
УСЛОВНА ВЕРОЯТНОСТ СЛЕД Е-ИЗМЕРВАНИЯ:(5)
ЕЛИМИНИРАНЕ НА ДИАГНОЗИ
КРАЙ НА ПРОЦЕСА
не
да
ИЗХОДНИ РЕЗУЛТАТИ
КРАЙ

Фигура 1.