Politechnika Śląska
Wydział Mechaniczny Technologiczny
Katedra Podstaw Konstrukcji Maszyn

Grzegorz Urbanek

Zastosowanie algorytmów ewolucyjnych w identyfikacji odwrotnych modeli diagnostycznych

Gliwice 2006

Recenzenci
Prof. dr hab. inż. Jan Kiciński, Instytut Maszyn Przepływowych PAN
Prof. dr hab. Wojciech Moczulski, Politechnika Śląska

Redaktor zeszytów
Wojciech Cholewa

Redaktor techniczny
Marek Wyleżoł

Projekt okładki
Wojciech Cholewa, Marek Wyleżoł

ISBN 83–60759–00–6

Wydawca
Katedra Podstaw Konstrukcji Maszyn
Wydział Mechaniczny Technologiczny
Politechnika Śląska
ul. Konarskiego 18a, 44-100 Gliwice
tel. (32) 237-14-67, fax (32) 237-13-60
https://kpkm.polsl.pl

Druk i oprawa (z dostarczonych materiałów)
Drukarnia Internetowa i-Print.com.pl, 44-100 Gliwice

Od autora

Zeszyt został opracowany na podstawie mojej rozprawy doktorskiej, wykonanej pod kierunkiem prof. dr. hab. inż. Wojciecha Cholewy. Publiczna obrona rozprawy odbyła się 22 czerwca 2006 roku przed Komisją powołaną przez Radę Wydziału Mechanicznego Technologicznego. W opracowaniu zostały uwzględnione uwagi recenzentów rozprawy doktorskiej: prof. dr. hab. inż. Jana Kicińskiego i prof. dr. hab. Wojciecha Moczulskiego.

Niniejszą książkę dedykuję mojej żonie Dominice dziękując za jej wyrozumiałość i cierpliwość. Składam serdeczne podziękowania Koleżankom i Kolegom z Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej za okazaną mi pomoc i życzliwość w czasie wykonywania pracy. Szczególnie zaś dziękuję prof. dr. hab. inż. Wojciechowi Cholewie za pomoc w realizacji badań oraz cenne uwagi krytyczne.

Gliwice, lipiec 2006 *Grzegorz Urbanek*

Część pracy finansowano ze środków budżetowych na naukę w latach 2001-2006, jako fragmenty projektów badawczych.

Spis treści

Od autora **3**

Wykaz ważniejszych oznaczeń **7**

1. Wstęp **9**

2. Algorytmy ewolucyjne **11**
2.1. Optymalizacja populacji za pomocą algorytmu ewolucyjnego 12
2.2. Reprezentacja potencjalnych rozwiązań zadania 13
2.3. Generowanie rozwiązania początkowego 14
2.4. Funkcja oceniająca 14
2.5. Selekcja i sukcesja 14
2.6. Operacje genetyczne 16
2.7. Warunki zakończenia działania algorytmu ewolucyjnego 18
2.8. Parametry algorytmu ewolucyjnego 18

3. Diagnozowanie maszyn **19**
3.1. Koncepcje procesu diagnozowania 19
3.2. Systemy doradcze 22
3.3. Pozyskiwanie relacji diagnostycznych 23
3.4. Modele obiektów 24
3.4.1. Modele strukturalne 24
3.4.2. Modele oparte na relacjach diagnostycznych 25
3.4.3. Koncepcja *czarnej skrzynki* 26
3.5. Modele diagnostyczne 27
3.5.1. Przekształcanie modeli 27
3.5.2. Modele odwrotne 28
3.5.3. Dekompozycja modelu 30
3.5.4. Zadanie identyfikacji modelu 32
3.5.5. Postać modelu 33
3.6. Metody identyfikacji modeli diagnostycznych 35
3.6.1. Aproksymacyjne modele odwzorowań 35
3.6.2. Interpolacyjne modele odwzorowań 35
3.6.3. Metody doskonalenia modeli interpolacyjnych 36

3.6.4. Klasyfikatory 37
3.7. Podsumowanie 41

4. Problem badawczy **43**
4.1. Cel pracy 44
4.2. Tezy pracy 44
4.3. Zakres pracy 45

5. Zastosowanie algorytmów ewolucyjnych w identyfikacji modeli odwrotnych **47**
5.1. Sterowanie badaniami symulacyjnymi z użyciem algorytmu ewolucyjnego . 47
5.2. Ewolucyjna optymalizacja zbioru przykładów wzorcowych 48
5.3. Ewolucyjna optymalizacja zbioru uwzględnianych cech 49
5.3.1. Identyfikacja modeli lokalnych 50
5.3.2. Wybór cech relewantnych 51
5.3.3. Podział w przestrzeni sygnałów dodatkowych 56
5.3.4. Klasyfikatory 56
5.3.5. Ocena jakości otrzymanych modeli 58

6. Przykład zastosowania opracowanej metody **59**
6.1. Obiekt badań 59
6.1.1. Dane uczące 60
6.2. Plan badań 61
6.2.1. Opracowane oprogramowanie 62
6.3. Przebieg badań 62
6.3.1. Dekompozycja modelu, wybór sygnałów dodatkowych 62
6.3.2. Podział w przestrzeni sygnałów dodatkowych 63
6.3.3. Wyznaczenie i ocena klasyfikatorów 74
6.4. Analiza wybranego wariantu 79
6.5. Podsumowanie 89

7. Podsumowanie i wnioski **91**
7.1. Podsumowanie 91
7.2. Wniosek ogólny 91
7.3. Wnioski szczegółowe 92
7.4. Kierunki dalszych badań 92

Bibliografia **93**

Streszczenie **103**

Wykaz ważniejszych oznaczeń

A	zbiór cech sygnałów dodatkowych
B	zbiór danych uczących
B_e	zbiór danych testujących
B_r	zbiór danych trenujących
C	zbiór symptomów diagnostycznych
d	miara odległości
$D[t]$	zbiór potomków w t–tym kroku algorytmu ewolucyjnego
f	sygnał diagnostyczny
F	zbiór obserwowanych sygnałów
h	miara podobieństwa
M	model obiektu
N	model diagnostyczny, odwrotny model diagnostyczny
p_s	prawdopodobieństwo selekcji
$p_i[t]$	i–ty osobnik populacji w t–tym kroku algorytmu ewolucyjnego
$P[t]$	populacja w t–tym kroku algorytmu ewolucyjnego
$R[t]$	zbiór rodziców w t–tym kroku algorytmu ewolucyjnego
s	cecha stanu
S	zbiór cech stanu obiektu
t	czas 'makro', licznik kroków algorytmu ewolucyjnego
τ	czas 'mikro'
U	obserwowany układ
v	cecha sygnału diagnostycznego
V	zbiór cech sygnałów diagnostycznych
X	zbiór wejść obiektu
Y	zbiór wyjść obiektu
Z	zbiór oddziaływań zewnętrznych

Rozdział 1

Wstęp

Diagnostyka techniczna jest dziedziną wiedzy, której celem jest określanie szeroko rozumianego stanu technicznego obiektów (środków technicznych) [17]. Jej powstanie wiąże się ze wzrostem złożoności środków technicznych oraz wzrostem odpowiedzialności realizowanych przez nie zadań. Początkowo, od lat pięćdziesiątych ubiegłego stulecia, rozwijała się jako *diagnostyka maszyn*. Lata siedemdziesiąte ubiegłego wieku przyniosły powstanie drugiej gałęzi diagnostyki technicznej — *diagnostyki procesów*.

Ze względu na zainteresowania autora, które skupiają się wokół diagnostyki maszyn, praca związana jest z tą właśnie gałęzią szeroko rozumianej diagnostyki technicznej. Cel diagnostyki maszyn można scharakteryzować poprzez realizowane przez nią zadania [16]:

- określenie bieżącego stanu technicznego (diagnoza),
- określenie przyczyn zaistnienia obecnego stanu (geneza),
- określenie horyzontu czasowego przyszłej zmiany stanu technicznego (prognoza).

W ostatnim czasie obserwuje się wzrost liczby zastosowań metod diagnostyki technicznej dla potrzeb przemysłowych pod postacią szeroko rozumianych systemów nadzoru diagnostycznego. Systemy te pozwalają na diagnozowanie obiektów wraz z sygnalizacją oraz graficzną prezentacją stanu lub zmian stanu obiektu. Ze względu na zastosowanie można wyróżnić systemy nadzoru diagnostycznego procesów przemysłowych [105] oraz systemy nadzoru diagnostycznego maszyn. W wielu przypadkach mamy do czynienia ze współdziałaniem obydwu rodzajów systemów diagnostycznych, przykładem mogą tu być rozwiązania systemów nadzoru diagnostycznego procesu produkcji energii elektrycznej [7, 34].

Rozwój techniki komputerowej powoduje coraz częstsze stosowanie metod sztucznej inteligencji do wspomagania wnioskowania diagnostycznego. Szczególną rolę odgrywają tu systemy doradcze (np. [21, 26, 37, 106]). Zastosowanie ich w nowoczesnych komputerowych systemach nadzoru diagnostycznego [20, 23, 29] wymaga opracowania skutecznych algorytmów pozwalających na identyfikację relacji diagnostycznych, będących podstawą określania reguł wchodzących w skład ich bazy wiedzy [32, 33]. Reguły takie wskazują możliwe przyczyny (np. niewyważenie wirnika) obserwowanych skutków (np. wysoka amplituda drgań podpory łożyskowej).

Poszukiwanymi relacjami diagnostycznymi przekształcającymi symptomy diagnostyczne w stany techniczne (relacje *symptom–stan*) mogą być modele odwrotne, czyli modele diagnostyczne, które można uzyskać m.in. w wyniku odwracania modeli symulacyjnych (modeli diagnozowanych obiektów). Ponieważ analityczne odwracanie złożonych modeli obiektów jest trudne (często niemożliwe) modele odwrotne identyfikuje się na podstawie zbiorów przykładów opisujących cechy stanu diagnozowanego obiektu oraz odpowiadające im cechy sygnałów diagnostycznych. Podstawowym zadaniem optymalizacji tak identyfikowanych modeli jest zadanie optymalizacji stosowanych zbiorów przykładów. Algorytmami optymalizacji o szerokich możliwościach adaptacji do różnych klas zadań są algorytmy ewolucyjne.

Powszechnie stosowaną w przemyśle klasą maszyn są maszyny wirnikowe. Bardzo często spełniają one ważne czy nawet krytyczne zadania. O ich zasadniczym znaczeniu świadczy duża liczba prac poświęconych diagnostyce tej klasy maszyn (np. [6, 33, 35, 40, 42, 53, 68, 72, 73, 80, 81, 92, 98, 109, 116, 118]). Wiele z tych prac dotyczy bardzo ważnej grupy maszyn jaką są turbiny energetyczne. Mając to na uwadze, w niniejszej pracy, skoncentrowano się na tej klasie maszyn. Na wybór ten miał również wpływ udział autora w pracach badawczych związanych z opracowaniem systemu nadzoru diagnostycznego turbozespołów energetycznych.

W pracy pokazano wyniki badań, których celem było wykazanie przydatności algorytmów ewolucyjnych w identyfikacji odwrotnych modeli diagnostycznych. Omówiono zagadnienia diagnozowania maszyn, wskazano na dużą przydatność modeli odwrotnych. Przedstawiono metody identyfikacji odwrotnych modeli diagnostycznych i związane z tym trudności. Pokazano możliwości omijania tych trudności dzięki zastosowaniu algorytmów ewolucyjnych. Zastosowanie zaproponowanej w pracy metody identyfikacji odwrotnych modeli diagnostycznych z użyciem algorytmów ewolucyjnych przedstawiono na przykładzie rozpoznawania niesprawności turbogeneratora energetycznego w zakresie przemieszczeń podpór łożyskowych.

Rozdział 2

Algorytmy ewolucyjne

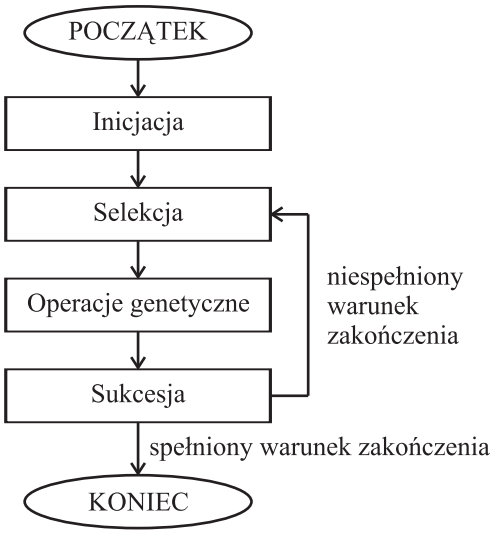
Algorytmy ewolucyjne są algorytmami heurystycznego przeszukiwania przestrzeni potencjalnych rozwiązań w celu znalezienia rozwiązania optymalnego. Sposób działania algorytmu ewolucyjnego jest wzorowany na zjawisku ewolucji biologicznej [57].

Algorytm ewolucyjny przetwarza populację osobników, z których każdy jest potencjalnym rozwiązaniem rozpatrywanego zadania (m.in.: [2, 54, 61, 89, 91, 100, 108]). Z każdym osobnikiem związany jest jego *genotyp* oraz *fenotyp*. Potencjalnym rozwiązaniem jest fenotyp osobnika, natomiast genotyp jest zakodowanym przepisem na utworzenie fenotypu. Genotyp opisany jest przez *chromosomy* (najczęściej jeden), które składają się z elementarnych części, zwanych *genami*. Pierwszym krokiem algorytmu ewolucyjnego jest utworzenie populacji początkowej — operacja ta zwana jest *inicjacją* (rys. 2.1). Na genotypach osobników wybranych w wyniku *selekcji* przeprowadzane są *operacje genetyczne*: *krzyżowanie* i *mutacja*. W ten sposób uzyskuje się genotypy nowych osobników, na których podstawie wyznacza się fenotypy. Funkcja oceniająca (*funkcja przystosowania*) ocenia osobnika na podstawie jego fenotypu. Wybór osobników, które utworzą populację do kolejnego kroku algorytmu ewolucyjnego nazywany jest *sukcesją*. Celem algorytmu ewolucyjnego jest optymalizacja populacji osobników, gdzie kryterium optymalizacji jest maksymalizacja funkcji przystosowania.

Algorytmy ewolucyjne, począwszy od lat 60. ubiegłego wieku, były przez długi czas rozwijane niezależnie w wielu ośrodkach naukowych na świecie; najważniejsze rodzaje algorytmów ewolucyjnych to [2]:

- prosty algorytm genetyczny [61, 64],
- strategie ewolucyjne [8, 102],
- programowanie ewolucyjne [9, 56],
- programowanie genetyczne [83, 84].

Zwiększone zainteresowanie algorytmami ewolucyjnymi w latach 90. ubiegłego wieku przyniosło wzajemne upodabnianie się ich odmian. Wtedy też, aby uniknąć problemów z rozstrzyganiem, do którego rodzaju rodzaju zaklasyfikować nowy pomysł czy modyfikację, ustalono wspólną nazwę: algorytmy ewolucyjne.



Rys. 2.1. Schemat algorytmu ewolucyjnego

Algorytmy ewolucyjne znalazły zastosowanie w różnych dziedzinach, m.in.:

- w procesie projektowo–konstrukcyjnym do optymalizacji postaci elementów maszyn (np. [11, 12, 62, 90])
- w projektowaniu układów elektronicznych (np. [86] (cyt. za [2]))
- w projektowaniu sieci telekomunikacyjnych (np. [85] (cyt. za [2]))
- w sterowaniu procesami przemysłowymi (np. [59] (cyt. za [60]))
- w diagnostyce technicznej (np. [18, 77, 78, 79, 82])

2.1. Optymalizacja populacji za pomocą algorytmu ewolucyjnego

Sformalizowany zapis algorytmu ewolucyjnego przedstawia algorytm 2.1.

Przeprowadzenie optymalizacji za pomocą algorytmu ewolucyjnego wymaga określenia:

- sposobu kodowania fenotypu w genotypie (sposobu reprezentacji potencjalnych rozwiązań w przestrzeni genotypu),
- sposobu tworzenia populacji początkowej,
- funkcji oceniającej (funkcji przystosowania),
- sposobu selekcji i sukcesji osobników,
- sposobu prowadzenia operacji genetycznych,

Algorytm 2.1 Podstawowy algorytm ewolucyjny; oznaczenia: t – czas (licznik kroków), P – populacja, R – rodzice, D – dzieci

procedura Algorytm_Ewolucyjny
{
  $t := 0$
  $P[t]$:=*Inicjacja*()
  Ocena($P[t]$)
  dopóki($\sim$ Warunek_zakończenia)
  {
    $t := t + 1$
    $R[t]$:=*Selekcja*($P[t-1]$)
    $D[t]$:=*Krzyżowanie_i_mutacja*($R[t]$)
    Ocena($D[t]$)
    $P[t]$:=*Sukcesja*($P[t-1], D[t]$)
  }
}

- warunku (warunków) zakończenia działania algorytmu,
- wartości różnych parametrów używanych w algorytmie ewolucyjnym (rozmiaru populacji, prawdopodobieństwa użycia operatorów genetycznych itp.).

2.2. Reprezentacja potencjalnych rozwiązań zadania

Wybór właściwej reprezentacji potencjalnych rozwiązań rozpatrywanego zadania wymaga rozpatrzenia: po pierwsze — co powinno stanowić fenotyp a co genotyp osobnika, po drugie — jakiego sposobu zapisu genotypu użyć. Przykładowo, jeżeli rozpatrywanym zadaniem byłoby poszukiwanie ekstremum funkcji to wartości argumentów tej funkcji będą genami składającymi się na chromosom będący genotypem osobnika, fenotypem natomiast będzie odpowiadająca tym wartościom argumentów wartość rozpatrywanej funkcji.

Sposoby zapisu chromosomów można podzielić na dwie grupy [25]:

1. chromosomy z liniowo ułożonymi genami, które mogą być kodowane binarnie, liczbowo lub w inny specjalny sposób,
2. chromosomy będące złożonymi strukturami genów (np. drzewa), tutaj geny mogą również być kodowane binarnie, liczbowo lub w inny specjalny sposób.

Powszechnie w algorytmach ewolucyjnych stosowane są chromosomy z liniowo ułożonymi genami kodowanymi binarnie. Wartości wszystkich genów każdego osobnika zostają przedstawione za pomocą liczb dwójkowych; „sklejone" razem stanowią chromosom. Reprezentacja ta ułatwia analizę teoretyczną i umożliwia wprowadzenie bardzo prostych operatorów genetycznych.

Innym sposobem jest zapis wartości genów w postaci dziesiętnej. Chromosom ma wtedy postać macierzy liczb, każda liczba jest wartością jednego z rozpatrywanych genów [58]. Taki zapis eliminuje konieczność przekształcania wartości genów do zapisu binarnego. Stosowanie takich chromosomów wymaga opracowania odpowiednich operatorów genetycznych.

2.3. Generowanie rozwiązania początkowego

Populacja początkowa najczęściej generowana jest w sposób losowy (o ile nie ma uzasadnionych podstaw, żeby postąpić inaczej). W przypadku kodowania binarnego z liniowo ułożonymi genami losuje się wartość każdego bitu chromosomu (łańcucha binarnego) i w ten sposób tworzy się całą populację. W przypadku kodowania liczbowego, wartość każdego genu jest losowana z przedziału zmienności wartości cechy, którą on reprezentuje.

2.4. Funkcja oceniająca

Funkcja oceniająca f_{oc} (funkcja przystosowania) ocenia jakość osobników (ich stopień przystosowania do przeżycia w środowisku). Funkcja ta jest jednym z najważniejszych elementów warunkujących skuteczne działanie algorytmu genetycznego. Nieodpowiednia funkcja oceniająca prowadzi do błędnego działania algorytmu. Niestety, nie istnieje ogólny „przepis" na opracowanie dobrej funkcji oceniającej. Funkcję taką zwykle wyznacza się metodą „prób i błędów" indywidualnie do konkretnego zadania.

2.5. Selekcja i sukcesja

Selekcja oraz sukcesja służą ukierunkowaniu algorytmu ewolucyjnego w stronę rozwiązań poprawiających dotychczasowe. Zadaniem selekcji (zwanej czasem preselekcją) jest wybór rodziców $R[t]$ czyli osobników, którzy zostaną poddani operacjom genetycznym. Sukcesja [2] (nazywana również postselekcją) służy utworzeniu populacji bazowej $P[t]$ do kolejnego kroku algorytmu na podstawie populacji potomnej (dzieci) $D[t]$ powstałej w wyniku operacji genetycznych oraz populacji $P[t-1]$ z poprzedniego kroku algorytmu.

Przykłady metod selekcji:

- *selekcja proporcjonalna* [2, 100], zwana również metodą ruletki (nazwa pochodzi od najprostszej realizacji algorytmicznej tego operatora [61]), jest procesem, w którym

chromosomy zostają powielone z prawdopodobieństwem wprost proporcjonalnym do wartości, jakie przybiera dla nich funkcja oceniająca; stosowanie tej metody jest możliwe tylko wtedy, gdy funkcja oceniająca przyjmuje wyłącznie wartości dodatnie; prawdopodobieństwo p_s selekcji osobnika $p_i[t-1]$ z populacji bazowej $P[t-1]$ do populacji rodziców $R[t]$ wyliczane jest z zależności:

$$p_s(p_i[t-1]) = \frac{f_{oc}(p_i[t-1])}{\sum f_{oc}(p_i[t-1])}, \tag{2.1}$$

- *selekcja turniejowa* [10] (cyt. za [89]) polega na powtarzaniu dwóch kroków: losowania par osobników wg metody ruletki, po wylosowaniu pary, osobnik o wyższym przystosowaniu zostaje ogłoszony zwycięzcą i umieszczony w populacji rodziców; proces jest kontynuowany aż do całkowitego wypełnienia tej populacji,

- *selekcja rangowa* [5] jest modyfikacją metody ruletki polegającą na wyznaczeniu prawdopodobieństwa selekcji osobnika na podstawie jego *rangi* $r(p_i[t-1])$ — liczby charakteryzującej jego jakość na tle populacji; liczbę tę można nadać następująco: sortuje się populację według nierosnących wartości oceny osobników, ranga jest tożsama z numerem porządkowym osobnika w tym uszeregowaniu; prawdopodobieństwo wyznacza się wtedy z zależności:

$$p_s(p_i[t-1]) = a + k(1 - \frac{r(p_i[t-1])}{r_{max}}), \tag{2.2}$$

  gdzie parametry a i k dobiera się tak, by były spełnione warunki:

$$\sum_{i=1}^{n} p_s(p_i[t-1]) = 1, \tag{2.3}$$

$$0 \leqslant p_s(p_i[t-1]) \leqslant 1, \tag{2.4}$$

- *selekcja progowa* [2, 14] jest odmianą selekcji rangowej, w metodzie tej w selekcji biorą udział tylko osobniki spełniające warunek:

$$0 \leqslant r(p_i[t-1]) < p_g, \tag{2.5}$$

  gdzie p_g jest zadanym progiem reprodukcji.

Przykłady metody sukcesji:

- *sukcesja trywialna* [2] polega na tym, że nową populacją staje się zbiór dzieci: $P[t] = D[t]$,

- *sukcesja elitarna* [2, 58] przeprowadzana jest w dwóch krokach: najpierw wybiera się z $P[t-1]$ k najlepszych osobników i przepisuje do $P[t]$, następnie na brakujące $m = n-k$ osobników wybiera się m najlepszych osobników ze złączenia $P[t-1] \cup D[t]$.

Ważnym zagadnieniem związanym z selekcją jest tzw. *nacisk selektywny*, który określa tendencję algorytmu do poprawiania średniej wartości przystosowania populacji bazowej [2]. Algorytm ma tym większy nacisk selektywny, im większa jest wartość oczekiwana liczby kopii lepszego osobnika w porównaniu z wartością oczekiwaną liczby kopii osobnika gorszego. W selekcji rangowej sterowanie naciskiem selektywnym odbywa się poprzez odpowiedni dobór rang, w selekcji progowej — przez właściwy dobór wartości progowej. Sterowanie naciskiem selektywnym w przypadku selekcji proporcjonalnej oraz turniejowej realizuje się poprzez *skalowanie przystosowania* [4, 61] czyli modyfikowanie szerokości przedziału wartości przyjmowanych przez funkcję przystosowania.

2.6. Operacje genetyczne

Populacja potomna $D[t]$ powstaje na skutek przeprowadzenia operacji genetycznych krzyżowania i mutacji na populacji rodzicielskiej $R[t]$. Operacje te polegają na losowych modyfikacjach genotypów osobników populacji rodzicielskiej. Operator krzyżowania działa na wielu (co najmniej dwóch) osobnikach populacji rodziców, efektem jego działania jest powstanie jednego lub wielu osobników potomnych, których chromosomy powstały w wyniku „wymieszania" odpowiednich genów chromosomów osobników rodzicielskich. Mutacja natomiast polega na modyfikacji genotypu jednego osobnika.

Najczęściej przyjmuje się, że w krzyżowaniu biorą udział dwa osobniki rodzicielskie, w wyniku otrzymuje się jeden lub dwa osobniki potomne. Zwykle osobniki do krzyżowania dobierane są w sposób losowy, nie jest to jednak regułą. Spośród innych sposobów kojarzenia osobników rodzicielskich można wymienić [65] (cyt. za [61]):

- kojarzenie wg linii — najlepszy osobnik jest kojarzony kolejno ze wszystkimi pozostałymi osobnikami,
- kojarzenie selektywne dodatnie polega na kojarzeniu osobników podobnych,
- kojarzenie selektywne ujemne to kojarzenie osobników niepodobnych,
- samozapłodnienie czyli kojarzenie osobnika z nim samym.

Przykłady metod krzyżowania:

- *krzyżowanie proste* (jednopunktowe) [61] przebiega następująco: wybiera się w sposób losowy (z jednakowym prawdopodobieństwem) punkt krzyżowania k spośród $l-1$ początkowych pozycji w chromosomie, po czym zamienia się miejscami wszystkie znaki od pozycji $k+1$ do l włącznie w obu elementach pary, tworząc w ten sposób dwa nowe ciągi;
- *krzyżowanie dwupunktowe* [15] (cyt. za [61]) przebiega analogicznie jak krzyżowanie jednopunktowe, z tą różnicą, że wybiera się dwa punkty cięcia;
- *krzyżowanie wielopunktowe* [47] (cyt. za [89]) jest uogólnieniem krzyżowania dwupunktowego na wiele punktów cięcia;

- *krzyżowanie jednorodne* [89] prowadzi się następująco: dla każdego bitu pierwszego potomka decyduje się (z zadanym prawdopodobieństwem), od którego z rodziców pobierze się wartość, drugi potomek otrzymuje bit od pozostałego rodzica;

- *rekombinacja puli genów* [89, 101]: kilku rodziców tworzy jednego potomka w ten sposób, że kolejne geny potomka są wybierane losowo z puli genów wybranych rodziców;

- *krzyżowanie uśredniające* [1] jest operatorem przeznaczonym dla reprezentacji liczbowej; w metodzie tej wartość każdego genu chromosomu potomnego jest liczbą zawierającą się między największą i najmniejszą wartością genu chromosomów rodzicielskich. Jeśli krzyżowanie uśredniające przebiega zgodnie ze schematem, że z pary chromosomów rodzicielskich powstaje para potomnych, wówczas chromosomy potomne będą symetryczne względem środka odcinka łączącego chromosomy rodzicielskie.

Przykłady metod mutacji:

- *mutacja równomierna* [61, 89] zmienia jeden lub więcej bitów na przeciwny z zadanym prawdopodobieństwem;

- *mutacja nierównomierna* [89] uwzględnia wiek populacji: wraz ze wzrostem wieku populacji bity znajdujące się bardziej na prawo w sekwencji kodu chromosomu otrzymują większe prawdopodobieństwo mutacji, znajdujące się zaś na lewo mniejsze;

- *mutacja rzeczywistoliczbowa* [1] stosowana jest wtedy, gdy geny przyjmują wartości ze zbioru liczb rzeczywistych, i polega na perturbacji wartości genu przez dodanie liczby wygenerowanej w sposób losowy.

Utworzone w wyniku operacji genetycznych osobniki potomne mogą nie spełniać ograniczeń rozwiązywanego zadania, tzn. mogą reprezentować rozwiązania niedopuszczalne. Ograniczenia zadania można w algorytmie ewolucyjnym uwzględniać w następujący sposób [2, 89]:

- poprzez „wbudowanie" ograniczeń w metodę generowania populacji początkowej oraz operatory genetyczne, wtedy wszystkie osobniki, które wygeneruje algorytm ewolucyjny są rozwiązaniami dopuszczalnymi,

- poprzez *funkcję kary*, która obniża ocenę osobników niespełniających ograniczeń,

- poprzez *algorytm naprawy*, który koryguje osobniki będące rozwiązaniami niedopuszczalnymi.

2.7. Warunki zakończenia działania algorytmu ewolucyjnego

Określenie odpowiedniego kryterium zatrzymania algorytmu ewolucyjnego nie jest zadaniem trywialnym. Najczęściej stosuje się kryteria, które polegają na monitorowaniu wartości funkcji przystosowania najlepszych osobników w kolejnych pokoleniach [2]:

- *kryterium maksymalnego kosztu* przyjmuje, że algorytm kończy działanie jeżeli jego koszt przekroczy pewną ustaloną wartość; najczęściej stosowaną jest odmiana, w której przyjmuje się pewną maksymalną dopuszczalną liczbę pokoleń generowanych przez algorytm;

- *kryterium zadowalającego poziomu funkcji przystosowania* kończy działanie algorytmu jeżeli algorytm znajdzie rozwiązanie o wartości funkcji przystosowania określonej przez użytkownika; określenie zadowalającej wartości funkcji przystosowania wymaga dobrej znajomości jej charakteru, zawsze też istnieje niebezpieczeństwo, że algorytm będzie działał dowolnie długo co w praktyce jest nieakceptowalne, z tego względu zwykle łącznie z tym kryterium uwzględnia się również kryterium maksymalnego kosztu;

- *kryterium minimalnej szybkości poprawy* zatrzymuje działanie algorytmu jeżeli nie uda się otrzymać lepszego rozwiązania w zadanej liczbie kolejnych pokoleń.

2.8. Parametry algorytmu ewolucyjnego

Tradycyjnie algorytmy ewolucyjne, operują na populacji, której liczebność w czasie optymalizacji nie zmienia się. Stosowane są również algorytmy ewolucyjne ze zmienną liczebnością populacji. Algorytmy takie nie zawierają żadnej odmiany mechanizmu selekcji, zamiast tego wprowadza się pojęcie „wieku" chromosomu, co jest równoważne liczbie pokoleń, przez które chromosom pozostaje „żywy". Tak więc wiek chromosomu pozwala na szczególne definiowanie pojęcia selekcji i, ponieważ zależy on od dopasowania osobnika, wpływa na liczebność populacji w każdym kroku procesu [3, 87].

Doboru wartości prawdopodobieństw krzyżowania oraz mutacji najczęściej dokonuje się doświadczalnie ze względu na brak szczegółowych zaleceń. Zwykle prawdopodobieństwo krzyżowania zawiera się w przedziale $[0,25; 0,75]$, zaś prawdopodobieństwo mutacji w przedziale $[0,005; 0,05]$.

Rozdział 3

Diagnozowanie maszyn

Spośród trzech zadań diagnostyki technicznej maszyn, czyli diagnozy, genezy oraz prognozy, najważniejsze, zdaniem autora, jest zadanie diagnozowania. Jedynie poprawne określenie bieżącego stanu maszyny nadaje sens dalszym działaniom, czyli określeniu przyczyn zaistnienia obecnego stanu oraz określeniu horyzontu czasowego przyszłej zmiany stanu technicznego.

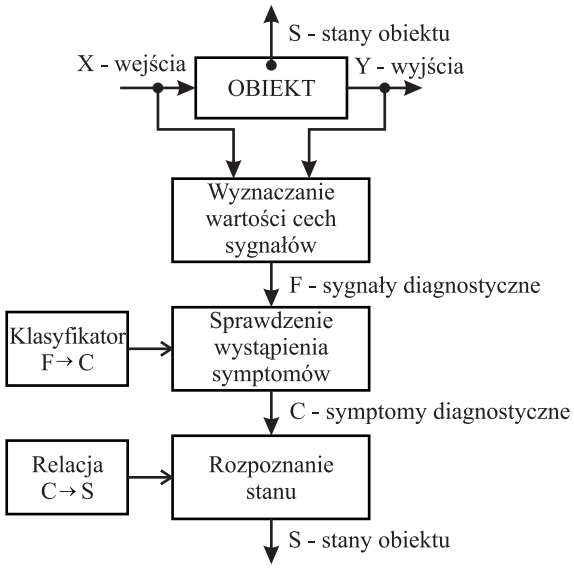
3.1. Koncepcje procesu diagnozowania

Na podstawie ogólnej koncepcji procesu diagnozowania można wskazać różnice pomiędzy diagnostyką maszyn a diagnostyką procesów. W diagnostyce procesów diagnozowanie przebiega w trzech fazach: detekcji, lokalizacji i identyfikacji uszkodzenia. W fazie detekcji na podstawie analizy sygnałów diagnostycznych wykrywane sa nieprawidłowości w przebiegu procesu. W fazie lokalizacji, na podstawie wykrytych symptomów, wskazywane są rodzaj i miejsce wystąpienia uszkodzenia. Faza identyfikacji uszkodzenia ma na celu ocenę stopnia zaawansowania uszkodzenia oraz podjęcie decyzji o przebiegu dalszego sterowania procesem. W diagnostyce maszyn wynikiem diagnozowania jest ocena stanu technicznego maszyny, która często sprowadza się do określenia klasy tego stanu. W związku z tym nie zawsze jest możliwe precyzyjne wyróżnienie trzech faz badań, fazy druga (lokalizacja) i trzecia (identyfikacja) najczęściej są nierozłączne.

W ramach diagnostyki maszyn można wyróżnić dwa nurty:

- diagnostykę symptomową,
- diagnostyką bazującą na modelach.

W diagnostyce symptomowej wnioskowanie o stanie obiektu prowadzone jest na podstawie cech sygnałów diagnostycznych (rys. 3.1). Dla obserwowanych sygnałów wejść i wyjść wyznacza się wartości ich cech. Następnie, na podstawie wyznaczonych wartości cech, sprawdza się wystąpienie symptomów określonych stanów z zastosowaniem odpowiedniego klasyfikatora. Rozpatrywane symptomy stanowią przesłanki reguł wnioskowania, na podstawie których następuje rozpoznanie stanu obiektu.



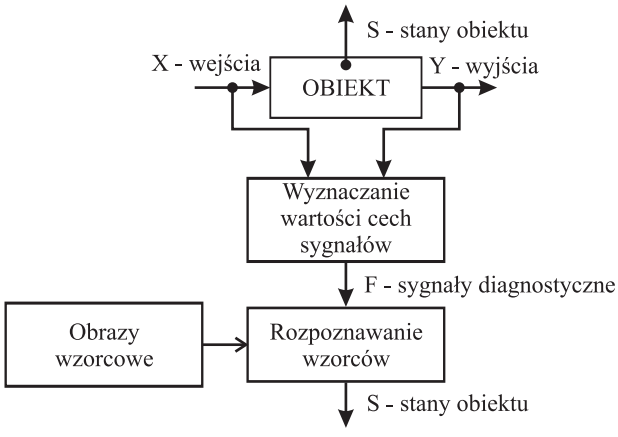
Rys. 3.1. Schemat procesu diagnozowania w diagnostyce symptomowej (na podstawie [36])

Diagnozowanie może być traktowane jako proces rozpoznawania wzorców (rys. 3.2), wówczas określenia stanu obiektu dokonuje się na podstawie wyznaczonych obrazów cech sygnałów diagnostycznych oraz obrazów wzorcowych.

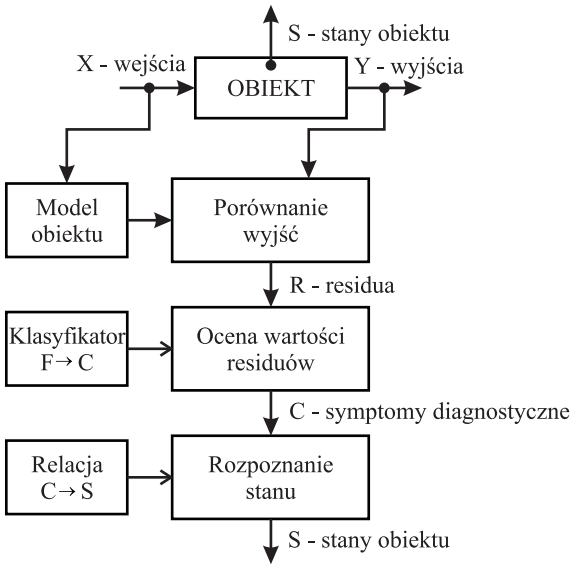
W diagnostyce bazującej na modelach można wyróżnić dwie klasy tych modeli:

- modele obiektów,
- modele diagnostyczne.

Schemat procesu diagnozowania z zastosowaniem modelu obiektu pokazano na rys. 3.3. Zakłada się, że model obiektu znajdującego się w określonym stanie technicznym (np. w stanie zdatnym) posiada wejścia i wyjścia odpowiadające wybranym wejściom i wyjściom obiektu, oraz że wyniki przekształcania wejść modelu w jego wyjścia są



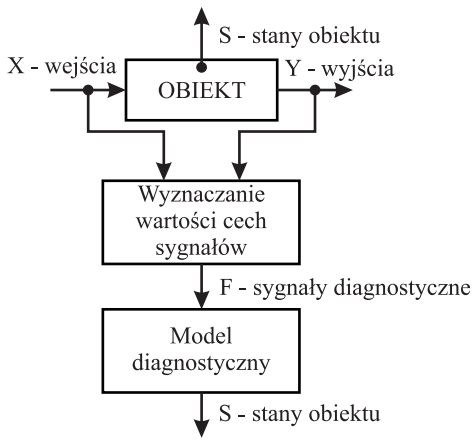
Rys. 3.2. Schemat procesu diagnozowania w oparciu o proces rozpoznawania wzorców (na podstawie [36])



Rys. 3.3. Schemat procesu diagnozowania z zastosowaniem modelu obiektu (na podstawie [36])

zgodne z wynikami takich przekształceń w obiekcie w rozpatrywanym stanie technicznym. W wyniku porównania sygnałów wyjściowych obiektu oraz sygnałów wyjściowych modelu generowany jest dodatkowy sygnał zwany residuum. Ocena wartości residuów jest podstawą określenia zbioru symptomów stanu, które pozwalają na określenie stanu obiektu.

Na rys. 3.4 pokazano schemat diagnozowania z zastosowaniem modelu diagnostycznego. Model diagnostyczny wyznacza cechy stanu obiektu na podstawie wartości wybranych cech sygnałów diagnostycznych.

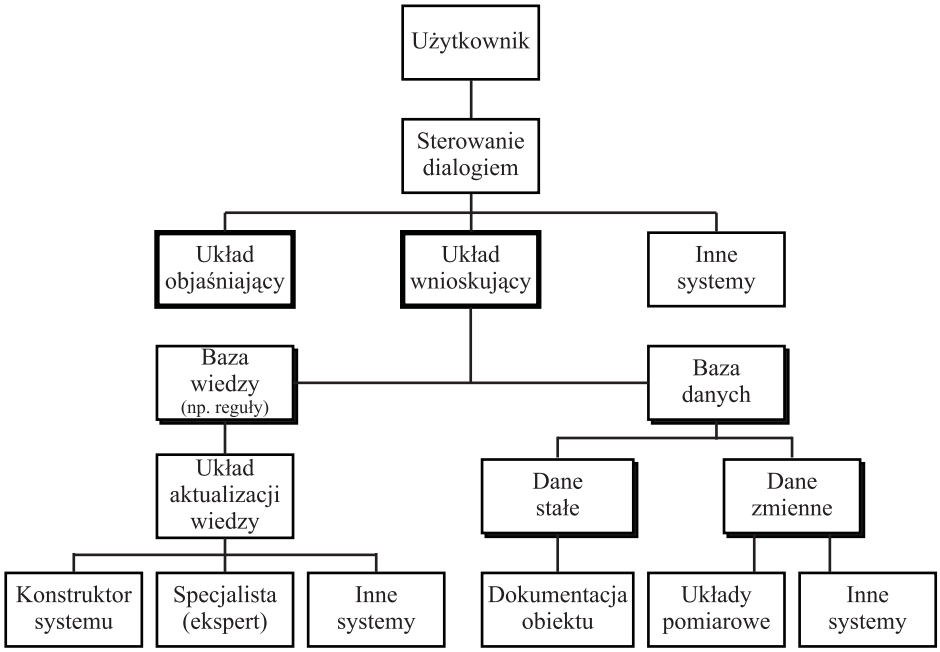


Rys. 3.4. Schemat procesu diagnozowania z zastosowaniem modelu diagnostycznego (na podstawie [36])

3.2. Systemy doradcze

Cechy sygnałów diagnostycznych, wyznaczone w wyniku analizy tych sygnałów, niosą dużą ilość informacji. Aby ułatwić proces podejmowania decyzji diagnostycznej, stosuje się dodatkowe środki wspomagające, oparte na metodach przetwarzania, redukcji i selekcji informacji. W przypadku złożonych systemów diagnozujących coraz częściej stosowane są rozwiązania oparte na systemach doradczych.

Systemy doradcze to programy komputerowe, przeznaczone do rozwiązywania specjalistycznych zadań, które wymagają profesjonalnej wiedzy i doświadczenia w stosowaniu tej wiedzy [24, 26, 94, 95]. Główne elementy systemu doradczego pokazano na rys. 3.5. Umożliwiają one stosowanie i aktualizację wiedzy pozyskanej w sposób jawny w wyniku analizy sygnałów i ich cech, a także komunikację z użytkownikiem.



Rys. 3.5. Główne elementy statycznego systemu doradczego [24]

Rozpatruje się dwie kategorie systemów doradczych: systemy statyczne, których bazy wiedzy nie zmieniają się w czasie działania systemu, oraz systemy dynamiczne, czyli systemy, których zawartość baz wiedzy zależy od czasu, systemy te posiadają możliwość adaptacji m.in. poprzez możliwość pobierania i uwzględniania informacji o trafności diagnozy. Niezależnie od tego podziału rozróżnia się systemy o nieograniczonym czasie działanie oraz systemy o ograniczonym i gwarantowanym czasie działania [24].

W procesie wspomagania wnioskowania diagnostycznego w systemach diagnostycznych maszyn wirnikowych szczególne znaczenie mają dynamiczne systemy doradcze [34]. Zastosowanie takich systemów w komputerowych systemach nadzoru diagnostycznego wymaga opracowania skutecznych algorytmów pozwalających na identyfikację relacji diagnostycznych, będących podstawą określania reguł wchodzących w skład baz wiedzy tych systemów.

3.3. Pozyskiwanie relacji diagnostycznych

Od wielu lat znane są w diagnostyce technicznej trzy klasy badań [16, 17, 121], pozwalających na identyfikację relacji diagnostycznych występujących między stanem technicznym a jego symptomami (relacje *stan–symptom*):

- *eksperymenty czynne*,
- *eksperymenty czynno–bierne*,
- *eksperymenty bierne*.

Eksperyment czynny polega na celowej zmianie cech stanu obiektu i obserwacji skutków jakie te zmiany powodują. Możliwość zmiany cech stanu jest warunkiem koniecznym tej formy badań eksperymentalnych; zmiana stanu oznacza jednak zwykle demontaż zespołów, co jest niestety często równoważne, po ponownym montażu, uzyskaniu innego stanu maszyny.

Tej formy eksperymentu nie można stosować w przypadku rzeczywistych dużych obiektów, takich jak np. turbogenerator, ze względu na zbyt duży koszt i stopień ryzyka. Dlatego też czynny eksperyment diagnostyczny stosowany jest najczęściej w warunkach laboratoryjnych dla potrzeb identyfikacji i optymalizacji modelu obiektu.

Eksperyment czynno–bierny polega na obserwacji sygnałów diagnostycznych z równoczesną obserwacją (pomiarem) cech stanu dla jednej (przy naprawie/awarii) lub wielu chwil (przy uruchomieniu i przy naprawie/awarii) czasu eksploatacji, bez możliwości ingerencji w wartości cech stanu.

W przypadku eksperymentu biernego nie ma możliwości ingerencji w budowę obiektu lub jego sterowanie. Obserwacji symptomów stanu nie towarzyszy ilościowa informacja o wartościach cech stanu. Dostępna może być jedynie informacja o klasie stanu.

Zaletą eksperymentu czynno–biernego i biernego jest niezakłócanie normalnej pracy obiektu co jest bardzo ważne w warunkach przemysłowych, wadą — długi czas zbierania informacji oraz ich fragmentaryczność ze względu na niewielki zazwyczaj zakres zmian stanu w czasie normalnej pracy. Szczególną wadą tej klasy eksperymentów jest często występujący w praktyce brak możliwości obserwowania obiektu w stanach krytycznych, które nie są osiągane przez poprawnie eksploatowany obiekt.

W ostatnim okresie nastąpił intensywny rozwój metod numerycznego modelowania obiektów technicznych i związanych z nimi procesów. Wynikiem modelowania są modele przyczynowo-skutkowe, które pozwalają na wyznaczenie skutków działania obiektu (np. drgań) dla złożonego stanu technicznego oraz dla zadanych cech konstrukcyjnych obiektu i warunków jego działania. Rozwinięto techniki strojenia modeli numerycznych, czyli techniki ich dopasowywania do wyników obserwacji działania obiektów rzeczywistych. Wysoka jakość odpowiednio strojonych modeli pozwala na przyjęcie założenia, że model jest wiernym odpowiednikiem badanego obiektu rzeczywistego, jego sygnały wejściowe i wyjściowe można utożsamiać z odpowiednimi sygnałami obiektu a uzyskane w wyniku jego badania wnioski można przenieść na rozważany obiekt rzeczywisty.

Brak możliwości szerszego stosowania eksperymentów czynnych na obiektach rzeczywistych oraz wysoka jakość modeli numerycznych pozwalają na zastępowanie eksperymentów czynnych na obiektach rzeczywistych *badaniami symulacyjnymi*, czyli eksperymentami czynnymi wykonywanymi na modelach symulatorów obiektów [22, 69]. Zaletą badań symulacyjnych jest możliwość ich przeprowadzenia nawet dla ekstremalnych warunków działania obiektu, niedopuszczalnych w czasie normalnej eksploatacji [70]. Prowadzenie takich badań wymaga jednak [28, 33]:

- dużej wiedzy dotyczącej parametrycznego doboru potencjalnych defektów w sposób umożliwiający otrzymanie określonego symptomu,
- odpowiedniej analizy prowadzonej przez osobę interpretującą wyniki badań symulacyjnych w celu przekształcenia tych wyników w relacje diagnostyczne.

3.4. Modele obiektów

Określając model obiektu (maszyny) należy rozróżniać dwa rodzaje takich modeli [31, 93]:

- modele indywidualne opisujące właściwości pojedynczego obiektu (określonego egzemplarza),
- modele grupowe opisujące łącznie właściwości określonej klasy obiektów (wszystkich egzemplarzy).

Najczęściej w rozważaniach diagnostycznych stosowane są modele grupowe, z tego względu dalsze rozważania będą dotyczyły tego typu modeli. Założenie to można uzasadnić tym, że modele indywidualne mogą być rozpatrywane jako szczególne przypadki modeli grupowych [31].

W dalszych rozważaniach szczególną uwagę zwrócono na zagadnienia związane z maszynami wirnikowymi, które są główną klasą obiektów rozpatrywanych w tej pracy.

Wiele klas i rodzajów modeli opisuje działanie i procesy zachodzące w maszynie wirnikowej [33, 35, 45, 76, 93, 95, 121]. Bardzo często stosowane są modele strukturalne odwzorowujące strukturę dynamiczną i wzajemne oddziaływania elementów maszyny [16, 63] oraz modele opisujące relacje między cechami stanu i cechami sygnałów diagnostycznych identyfikowane na podstawie eksperymentów [31, 93].

3.4.1. Modele strukturalne

Modele strukturalne umożliwiają wnioskowanie o rodzaju wielkości fizycznych, których zmiany powinny być obserwowane podczas badań jako sygnały diagnostyczne, o rodzaju cech tych sygnałów, które prawdopodobnie są wrażliwe na zmiany stanu obiektu, o zalecanej lokalizacji czujników pomiarowych oraz o wymaganych warunkach działania obiektu podczas badań diagnostycznych [19].

Modele strukturalne wykorzystywane są do analizy dynamiki wirnika czy też układu wirnik–podpory–fundament. Na podstawie wyznaczonych równań ruchu oraz przyjętych warunków brzegowych budowane są modele matematyczne. Budując każdy z modeli matematycznych, niezależnie od wzajemnego zróżnicowania, w równaniach ruchu uwzględnia się wpływ [63]:

- elementów bezwładnych (inercyjnych); są nimi z reguły tarcze i sprzęgła, natomiast bezwładność wału i podpór bywa pomijana,
- elementów sprężystych; są nimi z reguły wały, często pomija się natomiast podatność podpór,
- czynników warunkujących ruch precesyjny (wirowanie); powodowane są one z reguły przez wyrównoważenie wirnika, efekt giroskopowy tarcz, anizotropię sprężystą oraz tarcie wewnętrzne wału, siły hydrodynamiczne (podłoże hydrauliczne) i siły aerodynamiczne (cyrkulacyjne).

Modele matematyczne (ich implementacje numeryczne) znalazły szerokie zastosowanie w badaniach diagnostycznych maszyn wirnikowych [33, 71, 76, 80, 95], uzyskane za ich pomocą wyniki obliczeń często stanowią podstawę do:

- identyfikacji relacji między cechami stanu i cechami sygnałów diagnostycznych,
- budowy zbioru możliwych stanów obiektu,
- pozyskania sygnałów z dowolnych punktów pomiarowych.

Identyfikacja modeli strukturalnych prowadzona jest przy przyjęciu pewnych uproszczeń, które często mogą skutkować brakiem zgodności szczegółowych wniosków dotyczących relacji między cechami sygnałów diagnostycznych i cechami stanu obiektu, wyznaczonych na podstawie modeli, z wnioskami co do relacji wyznaczonych na podstawie eksperymentów [19].

3.4.2. Modele oparte na relacjach diagnostycznych

Modele oparte na relacjach diagnostycznych bazują na ogólnej teorii systemów. Pozwala to na jednolite rozpatrywanie obiektów opisanych za pomocą wejść, wyjść i stanów. Ważną cechą podejścia systemowego jest to, że do zidentyfikowania modelu, nie wystarcza specyfikacja elementów składowych, konieczne jest również zidentyfikowanie relacji zachodzących między tymi elementami (włącznie z otoczeniem). W badaniach diagnostycznych wspomaganych modelem obiektu można identyfikować następujące klasy relacji diagnostycznych (rys. 3.6) [33]:

a: relacje *wejście–wyjście* (model obiektu),

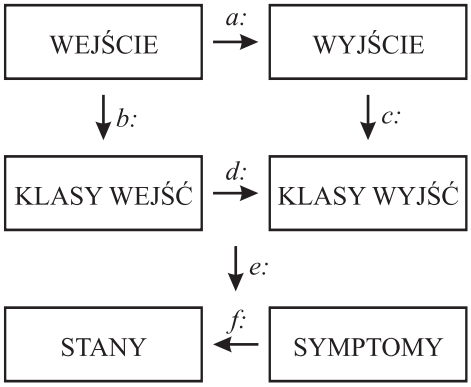
b: klasyfikator wejściowych cech modelu (nauczyciel) pozwalający na przekształcenie cech ilościowych w jakościowe,

c: klasyfikator wyjściowych cech modelu, pozwalający na przekształcenie cech ilościowych w jakościowe,

d: zmodyfikowany (przybliżony model obiektu),

e: operacja odwracania modelu,

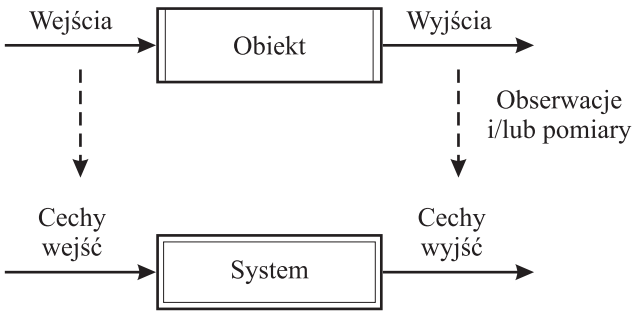
f: model odwrotny pozwalający na wnioskowanie o klasie stanu technicznego na podstawie symptomów.



Rys. 3.6. Relacje diagnostyczne [33]

3.4.3. Koncepcja czarnej skrzynki

W ujęciu systemowym [31, 50] obiekt traktowany jest jako *czarna skrzynka*. Pozwala to na pominięcie rozważań dotyczących struktury wewnętrznej obiektu. Wtedy obiekt rozpatrywany jest jako wyodrębniony myślowo z otoczenia poprzez określenie umownej granicy pomiędzy rozpatrywanym obiektem a jego otoczeniem. Oddziaływanie otoczenia na obiekt reprezentowane jest w postaci wejść obiektu, a oddziaływanie obiektu na otoczenie reprezentowane jest w postaci wyjść obiektu (rys. 3.7), przy czym zbiory wejść i wyjść są rozłączne.



Rys. 3.7. Model obiektu w ujęciu systemowym (na podstawie [31])

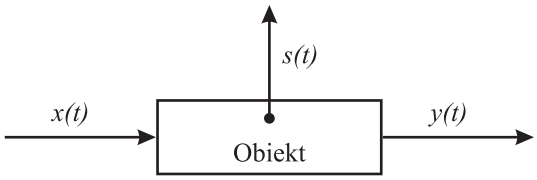
Obserwacja wejścia oraz wyjścia obiektu może się odbywać za pośrednictwem sygnałów wejściowych oraz sygnałów wyjściowych. Wyznaczone cechy sygnałów mogą być interpretowane jako cechy wejścia i wyjścia systemu.

Rozdzielenie wszystkich oddziaływań na rozłączne zbiory wejść i wyjść jest zadaniem trudnym, często niemożliwym do zrealizowania. Dlatego przyjmuje się domyślne założenie że model ma charakter przyczynowo-skutkowy, czyli każde wyjście modelu w dowolnej chwili czasu zależy wyłącznie od stanu wejść we wcześniejszych chwilach czasu i nie zależy od stanu wejść w przyszłości. Należy zwrócić uwagę, że warunek ten nie gwarantuje, że rozpatrywany model jest modelem przyczynowo-skutkowym. Dla systemu obiektu rzeczywistego uwzględnia się wszystkie możliwe (praktycznie nieskończenie wiele) wejścia i wyjścia, a dla systemu modelu uwzględnia się skończoną liczbę wejść i wyjść, niekompletny zbiór wejść często prowadzi do braku przyczyn wyjaśniających zmiany wyjść [31].

3.5. Modele diagnostyczne

3.5.1. Przekształcanie modeli

Systemy większości rzeczywistych obiektów technicznych są systemami dynamicznymi. W systemach takich cechy wyjść w danej chwili czasu zależą nie tylko od wartości cech wejść w tej samej chwili czasu ale również od cech wejść w chwilach poprzedzających. Historia systemu wynikająca ze stanu wejść do zadanej chwili czasu może być opisana za pomocą cech wewnętrznych obiektu. Cechy te są reprezentowane jako zbiór zmiennych losowych nazywanych zmiennymi stanu systemu (rys. 3.8) [31].



Rys. 3.8. Model obiektu: *x(t)* — sygnał wejściowy, *y(t)* — sygnał wyjściowy, *s(t)* — zmienna stanu [31]

Określenie przestrzeni cech wyjść jest możliwe na podstawie znajomości przestrzeni cech wejść oraz przestrzeni cech zmiennych stanu. Stanowi to podstawę budowy modelu obiektu. Model ten dla ustalonego czasu można zapisać w postaci:

$$Y = X \times S. \tag{3.1}$$

Dla wybranego obiektu w danej chwili czasu model (3.1) można zapisać w postaci relacji [31]:

$$M \subset X \times Y \times S. \tag{3.2}$$

gdzie:
M — model dla wybranego obiektu,
X — przestrzeń cech wejść,
Y — przestrzeń cech wyjść,
S — przestrzeń cech stanów.

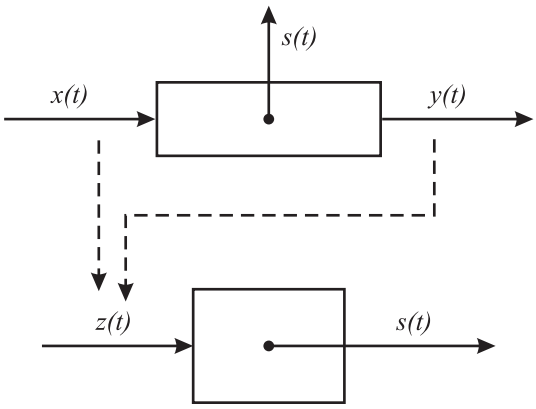
Głównym zadaniem badań diagnostycznych jest wykrywanie zmian stanu obiektu na podstawie cech oddziaływań zachodzących między obiektem i otoczeniem. Mając to na uwadze łączy się cechy wejść z cechemi wyjść (rys. 3.9) otrzymując w ten sposób cechy oddziaływań zewnętrznych. Pozwala to na zapisanie *modelu diagnostycznego* obiektu [31]:

$$N \subset Z \times S, \tag{3.3}$$

gdzie:

$$Z = X \times Y, \tag{3.4}$$

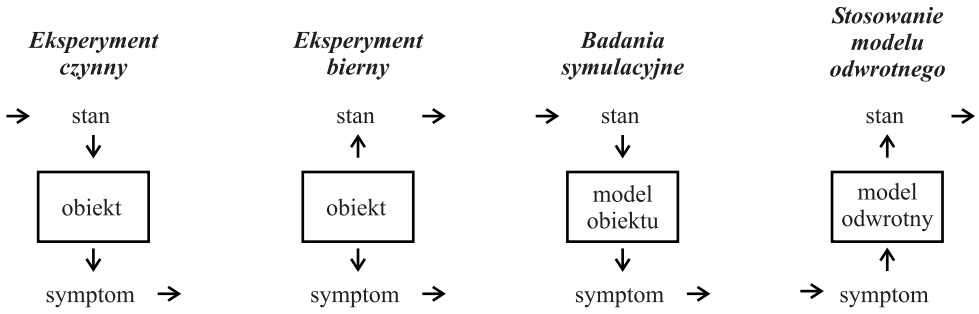
N — model diagnostyczny wybranego obiektu,
Z — przestrzeń cech oddziaływań zewnętrznych.



Rys. 3.9. Przekształcanie modeli [31]

3.5.2. Modele odwrotne

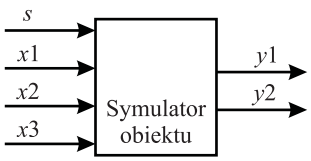
Odwrotny model diagnostyczny (rys. 3.10) jest modelem diagnostycznym otrzymanym w wyniku odwracania znanego modelu numerycznego rozpatrywanego obiektu [33, 42, 43, 44]. Znajomość modeli odwrotnych eliminuje potrzebę jawnego sterowania badaniami symulacyjnymi i umożliwiających pełną identyfikację reguł diagnostycznych, w których przesłanką jest symptom a konkluzją stan techniczny.



Rys. 3.10. Metody identyfikacji relacji diagnostycznych [33]

Zasadniczą trudność dla stosowania modeli odwrotnych stanowi brak możliwości analitycznego odwracania złożonych modeli (zwłaszcza nieliniowych). Problem ten wyeliminowano w [33] zakładając, że odwracanie modelu obiektu jest przeprowadzane na podstawie przykładów opisujących zadane stany rozpatrywanego obiektu oraz odpowiadające tym stanom symptomy.

W koncepcji poszukiwania modeli odwrotnych [42] zakłada się, że znany jest model symulatora obiektu (rys. 3.11), który pozwala na wyznaczenie wartości wyjść $\{y1, y2\}$ dla zadanego zbioru wartości wejść $\{x1, x2, x3\}$ i określonego stanu $\{s\}$.



Rys. 3.11. Wejścia i wyjścia modelu symulatora obiektu (na podstawie [42])

Model ten może być dany np. w postaci numerycznej. Wymienione wartości wejść i wyjść w zastosowaniach diagnostycznych mogą mieć następujące znaczenie [42]:

s wartości poszukiwane jako wynik procesu diagnozowania (np. rozkład niewyrównoważenia elementów wirujących),

$x1$ znane wartości, określające warunki działania obiektu (np. wartość prędkości obrotowej, wartość obciążenia),

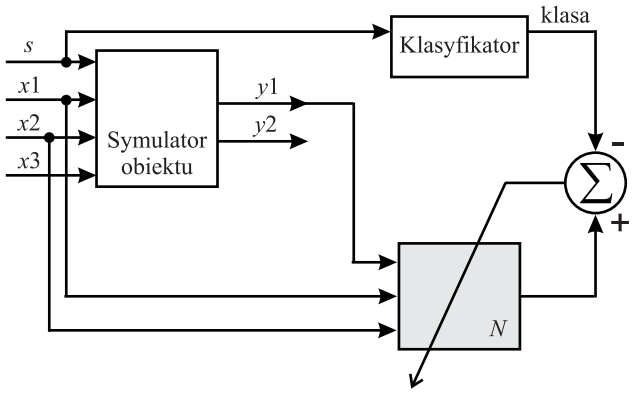
$x2$ znane wartości, które dla rozpatrywanego obiektu nie ulegają zmianom (np. wartości cech konstrukcyjnych obiektu),

$x3$ wartości trudne do oszacowania i pomiaru, które należy przyjmować jako wartości losowe wpływające na ograniczenie dokładności modelu (np. wartość tłumienia drgań przez posadowienie rozpatrywanego obiektu),

$y1$ znane wartości, będące skutkiem działania obiektu i określane w wyniku pomiaru lub odpowiedniej symulacji (np. wartości cech drgań),

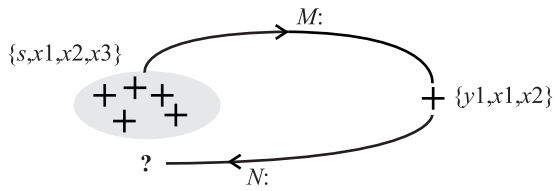
$y2$ nieznane lub pomijane wartości, będące skutkiem działania obiektu.

Model symulatora obiektu opisuje relacje przyczynowo-skutkowe występujące pomiędzy wartościami wejściowymi — przyczynami $\{s, x1, x2, x3\}$ oraz wartościami wyjściowymi — skutkami $\{y1, y2\}$. Odwracania modelu obiektu ma prowadzić do uzyskania modelu odwrotnego N (rys. 3.12), przekształcającego znane, określane w wyniku pomiaru lub symulacji wartości $\{y1, s1\}$ w poszukiwane wartości $\{s\}$. Należy podkreślić, że w kontekście relacji przyczynowo-skutkowych wartości $\{s, x1, x2, *\}$ opisują wyłącznie wybrane przyczyny, a wartości $\{y1, *\}$ wybrane ich skutki. Oznacza to, że można przypuszczać, iż dokładny model odwrotny nie istnieje, ponieważ brak jest podstaw do zakładania, iż model obiektu jest odwzorowaniem jedno-jednoznacznym. Wynikających



Rys. 3.12. Wyznaczanie modelu odwrotnego N dla danego modelu symulatora obiektu z zastosowaniem klasyfikatora (na podstawie [42])

stąd trudności (zadanie może nie posiadać rozwiązania — rys. 3.13) można uniknąć przez ograniczenie dokładności poszukiwanego modelu odwrotnego, zakładając, iż będzie wyznaczał on klasy wartości parametrów, a nie dokładne wartości tych parametrów [42].



Rys. 3.13. Zadanie odwracania modelu może nie posiadać rozwiązania [42]

3.5.3. Dekompozycja modelu

Identyfikacja odwrotnych modeli diagnostycznych, na podstawie przykładów uzyskanych z numerycznych modeli symulatorów obiektów, jest zadaniem trudnym ze względu na występowanie [28]:

- nadmiernej liczby możliwych do wyznaczenia cech sygnałów diagnostycznych,
- dużej liczby cech stanu, na podstawie których identyfikowana jest klasa stanu,

co znacznie zwiększa rozmiary zadania identyfikacji modelu odwrotnego.

Sposobem na ograniczenie rozmiarów zadania jest zaproponowana w [28] dekompozycja globalnego modelu obiektu na modele lokalne. Prosta dekompozycja przestrzenna na modele lokalne związane z zespołami, podzespołami czy elementami rozpatrywanego obiektu może prowadzić do zbioru sprzężonych ze sobą modeli lokalnych. W celu uzyskania ograniczenia liczby rozmiarów rozwiązywanego zadania, należy poszukiwać niezależne, rozprzęgnięte modele lokalne.

Zaproponowanym w [28] sposobem rozprzęgnięcia modeli lokalnych jest wprowadzenie cech sygnałów dodatkowych występujących w roli zmiennych instrumentalnych, które mogą dotyczyć oddziaływań wirtualnych lub ukrytych, przy czym nie jest wymagane

żeby dotyczyły one obserwowalnych oddziaływań rzeczywistych. Wybór sygnałów dodatkowych oraz ich cech nie jest zadaniem trywialnym. Wybór ten można przeprowadzić następującymi sposobami [27, 28]:

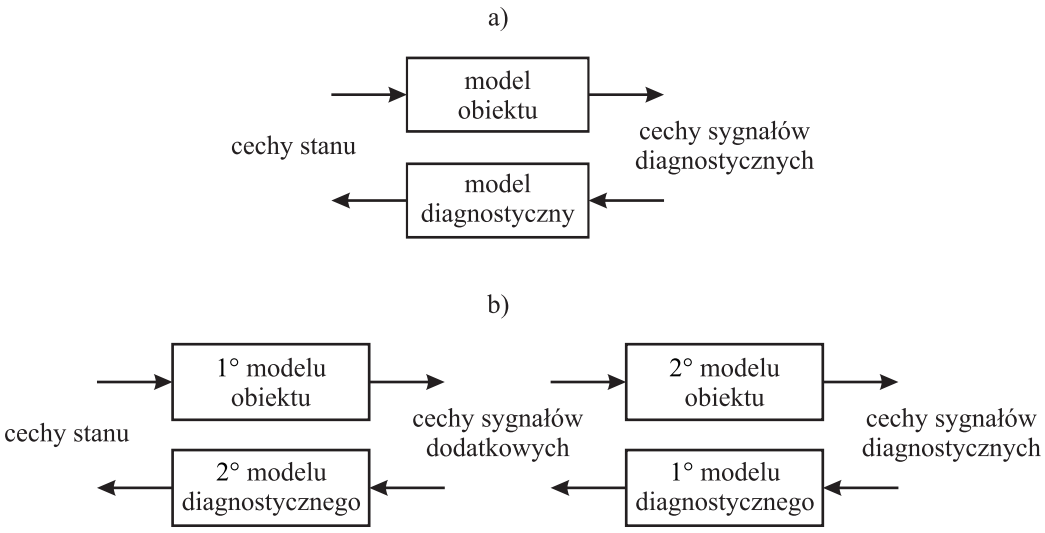
- poprzez poszukiwanie metodą prób i błędów (można tutaj zastosować algorytmy przeszukiwania heurystycznego, np. algorytmy ewolucyjne); metoda ta nie gwarantuje zadowalających wyników,
- korzystając z odpowiedniej wiedzy dziedzinowej, a w szczególności z wiedzy o funkcjonowaniu obiektu,
- wykorzystując wiedzę specjalistów; jest to dogodny sposób ze względu na to, że specjaliści potrafią wskazać sygnały i ich cechy, ważne ze względu na działania obiektu, dla których jednocześnie znane są modele opisujące ich związek z warunkami działanie i stanem technicznym obiektu.

Obiekt badań poddany dekompozycji może być rozpatrywany jako układ złożony z dwóch podukładów:

- układ pierwszego stopnia (1°) przekształcający wartości cech stanu w wartości cech sygnałów dodatkowych;
- układ drugiego stopnia (2°) przekształcający wartości cech sygnałów dodatkowych w wartości cech sygnałów obserwowanych (w symptomy).

Identyfikacja modelu odwrotnego tak zdekomponowanego obiektu może być realizowana jako identyfikacja modelu dwustopniowego (rys. 3.14) co wymaga rozwiązania następujących zadań cząstkowych:

- wyznaczenia pierwszego stopnia modelu diagnostycznego jako modelu odwrotnego dla wyróżnionego drugiego stopnia układu reprezentującego rozpatrywany obiekt,



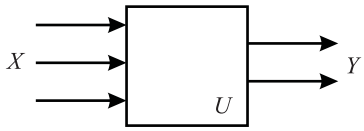
Rys. 3.14. Model obiektu i model diagnostyczny: a) model jednostopniowy, b) model dwustopniowy

- wyznaczenia drugiego stopnia modelu diagnostycznego jako modelu odwrotnego dla wyróżnionego pierwszego stopnia układu reprezentującego rozpatrywany obiekt,
- wyznaczenie łącznego modelu diagnostycznego przez szeregowe złożenie obu stopni modelu diagnostycznego, gdzie wyjście pierwszego stopnia jest wejściem drugiego stopnia.

Dużą zaletą takiego postępowania jest możliwość niezależnej identyfikacji każdego ze stopni modelu diagnostycznego.

3.5.4. Zadanie identyfikacji modelu

Zadanie identyfikacji modelu jest zadaniem bardzo ogólnym. Zakłada się, że poszukiwany model odwrotny (rys. 3.12) jest modelem jakiegoś wirtualnego, statycznego (nie zmieniającego się w czasie) układu U (rys. 3.15). Układ ten każdemu zbiorowi wartości wejść X przyporządkowuje odpowiedni zbiór wartości wyjść Y.



Rys. 3.15. Obserwowany wirtualny układ U (na podstawie [33])

Jeżeli postać modelu nie jest znana wtedy rozwiązanie zadania poszukiwania modelu wymaga wykonania dwóch kroków:

- wyboru postaci modelu,
- identyfikacji modelu (określenia wartości parametrów modelu o wybranej wcześniej postaci).

W celu określenia działanie układu U jest on obserwowany przez zewnętrznego obserwatora. Dla obserwatora tego dostępne są, w wyniku obserwacji w chwili t, następujące zbiory liczb [33]:

- zbiór wartości wyjść

$$\{y_1(t), y_2(t), ..., y_{L_y}(t)\}, \tag{3.5}$$

- zbiór wybranych wartości wejść

$$\{x_1(t), x_2(t), ..., x_{L_x}(t)\}, \tag{3.6}$$

takie, że dla każdej chwili czasu t

$$y_{l_y}(t) \in Y_{l_y}; l_y = 1, 2, ..., L_y, \tag{3.7}$$

$$x_{l_x}(t) \in X_{l_x}; l_x = 1, 2, ..., L_x, \tag{3.8}$$

gdzie:

$$Y_1, Y_2, ..., Y_{L_y}, \tag{3.9}$$

$$X_1, X_2, ..., X_{L_x}, \tag{3.10}$$

są przestrzeniami wartości wyjść i wejść.

Ogólna liczba wejść L_x oraz wartości pozostałych, nie objętych obserwacją, wartości wejść

$$\{x_{L_x+1}, x_{L_x+2}, ...\}, \tag{3.11}$$

nie są znane. Wartości wejść (3.11) traktuje się jako zakłócenia. Wynikiem obserwacji prowadzonych w chwilach

$$t \in \{t_1, t_2, ..., t_{L_t}\}, \tag{3.12}$$

są uporządkowane zbiory wartości

$$\begin{array}{rcl} B_1 & = & \{x_{1,1}, x_{1,2}, ..., x_{1,L_x}, y_{1,1}, y_{1,2}, ..., y_{1,L_y}\} \\ B_2 & = & \{x_{2,1}, x_{2,2}, ..., x_{2,L_x}, y_{2,1}, y_{2,2}, ..., y_{2,L_y}\} \\ \dots & & \\ B_{L_t} & = & \{x_{L_t,1}, x_{L_t,2}, ..., x_{L_t,L_x}, y_{L_t,1}, y_{L_t,2}, ..., y_{L_t,L_y}\} \end{array} \tag{3.13}$$

Zbiory wartości wyjść (3.5) oraz zbiory wartości wejść (3.6) mogą być zapisane w postaci macierzy

$$Y(t) = [y_1(t), y_2(t), ..., y_{L_y}(t)], \tag{3.14}$$

$$X(t) = [x_1(t), x_2(t), ..., x_{L_t}(t)]. \tag{3.15}$$

Macierze (3.14),(3.15) mogą być interpretowane jako współrzędne punktów przestrzeni wielowymiarowych. Kolejne współrzędne takiej przestrzeni odpowiadają rozpatrywanym wejściom i wyjściom badanego obiektu. Wartości współrzędnych są wartościami wejść i wyjść [33].

Poszukiwany model ma umożliwiać wyznaczanie zbioru wartości wyjść (3.5) dla zadanego zbioru wartości wejść (3.6). Brak znajomości struktury i działania układu U uniemożliwia uwzględnienia informacji o strukturze lub działaniu podczas wyboru postaci modelu opisującego układ U. Postać tego modelu należy dobrać odpowiednio do zbioru danych (3.13), zbiór ten nazywany jest zbiorem danych uczących.

3.5.5. Postać modelu

Dla tak określonego zadania istnieje wiele metod wyznaczania modeli. Modele te można podzielić na klasy według następujących kryteriów [28, 33]:

- rodzaj zależności stanowiących elementy modelu, które mogą być identyfikowane na podstawie zbiorów wartości wejść i wyjść:
  - *modele odwzorowań* zawierające funkcje odwzorowujące zbiory wartości wejść w zbiory wartości wyjść,

  - *modele relacji* zawierające relacje określające podzbiory, odpowiadających sobie wartości wejść i wartości wyjść, w iloczynie kartezjańskim przestrzeni wartości wejść i wyjść,

- istota działania modelu:
  - *modele aproksymacyjne*, zastępujące (z zadaną dokładnością) zbiór danych uczących,
  - *modele interpolacyjne*, rozpięte na zbiorach przykładów stanowiący odpowiedni podzbiór danych uczących i wymagające umieszczenia tego podzbioru w definicji modelu,

- stopień złożoności modelu:
  - *modele uogólniające*, w których liczba parametrów niezbędnych do zapisania modelu jest znacznie mniejsza od ogólnej liczby danych uczących, na podstawie których modele te zostały wyznaczone,
  - *modele pełne*, w których liczba parametrów niezbędnych do zapisania modelu jest porównywalna z liczbą danych uczących, na podstawie których model ten został wyznaczony,

- zakres modelu:
  - *modele globalne*, opisujące dane uczące w pełnej przestrzeni ich wartości,
  - *modele lokalne*, opisujące dane uczące w wybranym obszarze przestrzeni stanowiącym otoczenie określonego punktu,

- jakość rozwiązania:
  - *modele dokładne* pozwalające na wyznaczenie rozwiązań dokładnych,
  - *modele przybliżone* pozwalające na wyznaczenie wyłącznie rozwiązań przybliżonych lub rozmytych.

- sposób reprezentacji danych (3.5) i (3.6):
  - *modele ilościowe* dla danych zapisanych w postaci liczb,
  - *modele jakościowe* dla danych reprezentowanych w sposób przybliżony,
  - *modele rozmyte* dla danych reprezentowanych w postaci liczb lub zbiorów rozmytych,
  - *modele statystyczne* dla danych reprezentowanych w postaci łącznych rozkładów gęstości prawdopodobieństw ich wartości,
  - *modele przekonań* dla danych reprezentowanych w postaci rozkładów przekonań o ich wartościach, gdzie modele przekonań są odpowiednikami modeli statystycznych, w których miary prawdopodobieństw zastąpiono miarami subiektywnych przekonań,

– *klasyfikatory* umożliwiające wyznaczenie danych wyjściowych (3.5) reprezentowanych w postaci zbiorów, grup (skupień) lub klas.

3.6. Metody identyfikacji modeli diagnostycznych

3.6.1. Aproksymacyjne modele odwzorowań

Aproksymacyjne modele odwzorowań są typowym przykładem globalnych modeli odwrotnych. Ich wyznaczanie polega na identyfikacji, na podstawie zbioru danych uczących, funkcji przekształcającej z odpowiednią dokładnością wartości wejść w oceny wartości wyjść. Wymaga to przyjęcia założenia, że jedną z własności danych jest to, że poszukiwana funkcja istnieje.

Modele aproksymacyjne mogą być wyznaczane jako dokładne i przybliżone. Dla wyznaczenia modelu dokładnego konieczne jest przyjęcie odpowiedniego kryterium jakości przybliżenia danych uczących. Powszechnie stosowane są kryteria minimalnoodległościowe, najczęściej stosowane jest tzw. kryterium najmniejszych kwadratów.

Podstawową klasę dokładnych modeli aproksymacyjnych stanowią modele liniowe. Charakteryzują się prostymi algorytmami ich wyznaczania. Wyróżnia się również modele nieliniowe — z tej klasy modeli obecnie najczęściej stosowane są modele bazujące na koncepcji sztucznej sieci neuronowej.

Niedogodnością wymienionych modeli aproksymacyjnych jest brak prostych sposobów przekazania informacji o dokładności modelu. Wady tej pozbawione są przybliżone modele aproksymacyjne. Można wyróżnić dwie klasy takich modeli [33]: modele jednostronne i modele częściowo jednostronne.

3.6.2. Interpolacyjne modele odwzorowań

Modele interpolacyjne są przykładem lokalnych modeli odwrotnych. Pozwalają na wyznaczenie przybliżonej wartości funkcji dla dowolnego argumentu na podstawie znanych przykładów wartości tej funkcji.

Modele interpolacyjne można wyznaczać stosując m.in. następujące algorytmy [33, 35]:

- W11 — w algorytmie tym wartość dla dowolnego argumentu wyznaczana jest jako średnia ważona znanych wartości funkcji,
- W12 — wyznacza się lokalny model liniowy w bezpośrednim otoczeniu wyróżnionego punktu na podstawie jego najbliższych sąsiadów,
- W13 — algorytm ten jest zmodyfikowaną wersją algorytmu W12, modyfikacja ta polega na zwiększeniu wpływu bezpośrednich sąsiadów rozpatrywanego elementu poprzez wprowadzenie dodatkowej funkcji oceniającej stopień ważności danych,

- W14 — algorytm ten jest modyfikacją algorytmu W12, pozwalającą na uzyskanie przez rozpatrywany model własności nieliniowych.

3.6.3. Metody doskonalenia modeli interpolacyjnych

Lokalne modele interpolacyjne są rozpięte na przykładach zawartych w zbiorach danych wzorcowych (wymagają umieszczenia tego zbioru w definicji modelu; nie zastępują danych) co jest ich podstawową zaletą: nie przechowuje się modelu tylko zbiór danych, model wyznacza się w chwili zapotrzebowania na niego na podstawie aktualnych danych, a co za tym idzie rozszerzenie uwzględnianych przykładów wzorcowych nie stanowi problemu. Bezpośredni wpływ na jakość tak wyznaczanego modelu odwrotnego ma (oprócz wyboru funkcji interpolującej przykłady) jakość użytego zbioru przykładów, a dokładniej, odpowiednie wypełnienie przykładami przestrzeni symptomów, równomierne, bez „dziur".

Adaptacyjne sterowanie badaniami symulacyjnymi

W [75, 76] przedstawiono metodę doskonalenia zbioru przykładów wzorcowych (jakość tych przykładów decyduje o jakości modelu odwrotnego). Zakładając, że wyznaczane modele odwrotne są doskonalone przez odpowiednią aktualizację zbioru przykładów, wyodrębniono dwa etapy badań symulacyjnych:

- badania symulacyjne podstawowe, których celem jest otrzymanie podstawowego zbioru przykładów,
- badania symulacyjne uzupełniające, których celem jest uzupełnienie podstawowego zbioru przykładów nowymi przykładami.

W planowaniu podstawowych badań symulacyjnych można zastosować ogólnie znane techniki planowania eksperymentu, np. planowanie dwupoziomowe, trójpoziomowe, wielopoziomowe itd. (np. [67, 88, 97]). Celem uzupełniających badań symulacyjnych jest generowanie nowych przykładów, mających wypełnić obszary przestrzeni cech sygnałów diagnostycznych, w których jest najmniej danych. Wypełnianie obszarów o najmniejszej liczbie danych nowymi przykładami prowadzi do poprawy jakości wyznaczanego modelu odwrotnego. Działania zmierzające do uzyskania nowych przykładów podzielono na cztery etapy:

1. identyfikacja w przestrzeni cech sygnałów diagnostycznych takiego obszaru o zadanym rozmiarze (np. objętości), który zawiera najmniej punktów oraz przyjęcie jego reprezentanta v_s;
2. wyznaczenie wartości parametrów lokalnego modelu odwrotnego;
3. zastosowanie określonego wcześniej modelu lokalnego i wyznaczenie za jego pomocą wartości cech stanów s_i warunkujących (wg doskonalonego modelu odwrotnego) takie cech sygnałów diagnostycznych v_s;

4. zastosowanie symulatora w celu wyznaczenia cech sygnałów diagnostycznych v_i odpowiadających stanowi s_i.

3.6.4. Klasyfikatory

Zadanie identyfikacji dokładnego modelu odwrotnego może nie posiadać rozwiązania (rys. 3.13). Rozwiązaniem tego problemu jest ograniczenie dokładności poszukiwanego modelu odwrotnego tak by wyznaczał klasy wartości cech stanu, a nie dokładne wartości tych cech (rys. 3.12). Wobec tego model odwrotny może być identyfikowany jako klasyfikator realizujący zadanie klasyfikacji przykładów do klas wartości cech stanu na podstawie wartości wybranych cech sygnałów diagnostycznych.

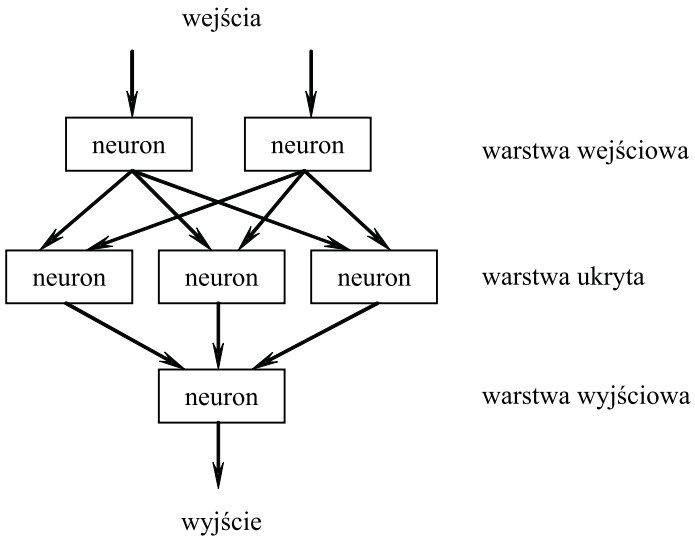
Znanych jest wiele metod i narzędzi klasyfikacji, do często stosowanych należą sztuczne sieci neuronowe, metoda wektorów wspierających, metody bazujące na miarach odległości i podobieństwa.

Sztuczne sieci neuronowe

Sztuczne sieci neuronowe stanowią próbę matematycznego odwzorowania działania ludzkiego mózgu oraz procesów w nim zachodzących.

Sztuczna sieć neuronowa składa się z połączonych ze sobą neuronów. W najczęściej stosowanych sieciach neurony ułożone są w tzw. warstwy (rys. 3.16) [48, 51, 107, 122]. Można wyróżnić:

- warstwę wejściową,
- warstwy ukryte,
- warstwę wyjściową.



Rys. 3.16. Warstwy sztucznej sieci neuronowej

Każdy neuron (rys. 3.17) przekształca swoje wejścia w wyjście wg zależności:

$$y = f(\Sigma(w_i x_i) + b), \tag{3.16}$$

gdzie:
x_i — wejścia neuronu,
y — wyjście neuronu,
f — funkcja przejścia (np.: liniowa, sigmoidalna, radialna),
w_i — wagi wejść,
b — składowa stała.

Rys. 3.17. Model neuronu

Projektując sztuczną sieć neuronową należy dla każdego neuronu określić początkowe wagi w_i wejść x_i, funkcję przejścia f oraz początkową wartość składowej stałej b.

Przed zastosowaniem sieci należy ją poddać trenowaniu, które polega na takim dobraniu wag w_i i wartości stałych b, by dla zadanych wartości wejściowych otrzymać żądane wartości wyjść, z minimalnym błędem.

Sztuczne sieci neuronowe znajdują zastosowanie głównie w zadaniach aproksymacji, niemniej jednak odpowiednie przygotowanie danych uczących pozwala na takie trenowanie sieci by spełniała zadanie klasyfikatora.

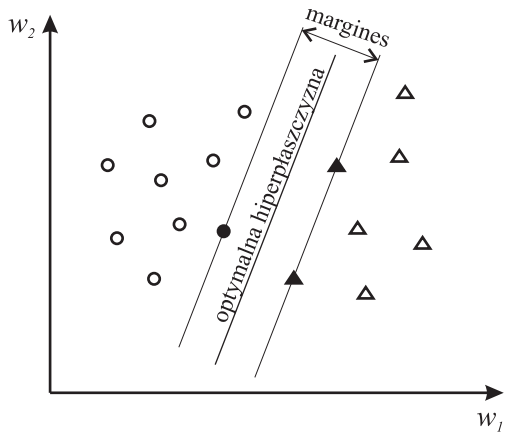
Metoda wektorów wspierających

Metoda wektorów wspierających (ang. Support Vector Machines — SVM) [13, 115, 120] jest metodą identyfikacji klasyfikatora i prowadzenia klasyfikacji polegającą na wyznaczeniu hiperpłaszczyzny separującej elementy należące do danej klasy od elementów nienależących do tej klasy.

Jeżeli zbiór B przykładów uczących zostanie określony następująco

$$B = \{(w_i, k_i)\}, w_i \in W \subset \mathbb{R}^l, k_i \in \{-1, 1\}, \tag{3.17}$$

gdzie:
w_i — wartości cech opisujących przykład,
k_i — informacja o przynależności ($k_i = 1$) lub braku przynależności ($k_i = -1$) do danej klasy,
l — liczba cech,
to identyfikacją klasyfikatora metodą SVM polega na poszukiwaniu możliwie najszerszego marginesu rozdzielającego przykłady należące i nienależące do rozpatrywanej klasy (rys. 3.18).



Rys. 3.18. Hiperpłaszczyzna separująca wektory należące (•) i nienależące (▲) do rozpatrywanej klasy

Pokazane na rys. 3.18 zadanie dotyczy danych należących do klas liniowo separowalnych. Dla zadań liniowo nieseparowalnych w przestrzeni cech $\mathbb{R}^l$ poszukuje się przestrzeni $\mathbb{R}^L$ o większej liczbie wymiarów ($L > l$), w której zadanie będzie liniowo separowalne. Przestrzeń ta jest określona funkcją jądra; do najczęściej stosowanych należą funkcje:

- liniowa
$$K(w_i, w_j) = w_i^T w_j, \tag{3.18}$$
- wielomianowa
$$K(w_i, w_j) = (w_i^T w_j + \theta)^p, \tag{3.19}$$
- radialna
$$K(w_i, w_j) = \exp\left(-\frac{\|w_i - w_j^2\|}{\sigma^2}\right). \tag{3.20}$$

Metody klasyfikacji bazujące na miarach odległości i podobieństwa

Klasyfikacja rozpatrywanego przykładu prowadzona jest na podstawie wybranej miary odległości czy też podobieństwa do klas zdefiniowanych poprzez reprezentantów tych klas (np. klasa może być zdefiniowana przez jednego reprezentanta).

Klasyfikację prowadzi się na podstawie przyjętego kryterium:

- najbliższego sąsiada — rozpatruje się po jednym reprezentancie każdej z klas, który jest najbliższym sąsiadem (wśród wszystkich reprezentantów danej klasy) klasyfikowanego przykładu; przykład zostaje zaklasyfikowany do klasy, do której należy najbliższy (najbardziej podobny) reprezentant,
- najdalszego sąsiada — analogicznie jak poprzednio, z tym że rozpatruje się najdalszych sąsiadów klasyfikowanego przykładu,
- elementu średniego — rozpatruje się reprezentantów będących najlepszym przybliżeniem środków klas; przykład zostaje zaklasyfikowany do klasy, do której należy najbliższy (najbardziej podobny) reprezentant,

- k najbliższych sąsiadów — rozpatruje się k najbliższych sąsiadów klasyfikowanego przykładu (wśród wszystkich reprezentantów klas), decyzję o klasyfikacji rozpatrywanego przykładu podejmuje się na podstawie ich przynależności do klas.

Prowadząc klasyfikację w l wymiarowej przestrzeni V wartości cech ilościowych można stosować następujące miary odległości pomiędzy elementami x, y tej przestrzeni [19]:

- Minkowskiego
$$d_1^p(x,y) = \frac{1}{l}\sum_{i=1}^{l}|x[i]-y[i]|^p, \tag{3.21}$$
- Euklidesa
$$d_2^2(x,y) = \frac{1}{l}\sum_{i=1}^{l}(x[i]-y[i])^2, \tag{3.22}$$
- Hamminga
$$d_3(x,y) = \frac{1}{l}\sum_{i=1}^{l}|x[i]-y[i]|, \tag{3.23}$$
- Sebestyena
$$d_4^2(x,y) = (x-y)W(x-y)^T, \tag{3.24}$$
gdzie:
W — diagonalna macierz wag:
$$W = \begin{bmatrix} w_1 & 0 & ... & 0 \\ 0 & w_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & w_m \end{bmatrix}, \tag{3.25}$$
- Mahalanobisa
$$d_5^2(x,y) = (x-y)C^{-1}(x-y)^T, \tag{3.26}$$
gdzie:
C^{-1} — odwrotna macierz macierzy kowariancji:
$$C = \frac{1}{l_e}\sum_{j=1}^{n}(v_j-\bar{v})^T(v_j-\bar{v}), v_j \in V, \tag{3.27}$$
l_e — liczba elementów rozpatrywanego zbioru w przestrzeni V,
$\bar{v}$ — średni element rozpatrywanego zbioru:
$$\bar{v} = \frac{1}{n}\sum_{j=1}^{n}v_j, \tag{3.28}$$
- Canberra
$$d_6(x,y) = \frac{1}{l}\sum_{i=1}^{l}\frac{|x[i]-y[i]|}{x[i]+y[i]}. \tag{3.29}$$

Dla określenia podobieństwa na podstawie funkcji odległości można zastosować różne funkcje podobieństwa. Przykładami takich funkcji są następujące zależności [19]:

$$h_1(x,y) = \frac{1}{1+\alpha d(x,y)^\beta}, \tag{3.30}$$

$$h_2(x,y) = 1 - \frac{d(x,y)}{\max_{(x,y)\in V\times V}(d(x,y))}, \tag{3.31}$$

$$h_3(x,y) = \exp(-\alpha d(x,y)), \tag{3.32}$$

gdzie:
α, β — stałe warunkujące własności funkcji h.

Funkcję podobieństwa można wyznaczyć bezpośrednio w postaci cosinusa kąta między wektorami wskazującymi elementy x, y przestrzeni V

$$h_4(x,y) = \frac{xy^T}{\sqrt{xx^T yy^T}}. \tag{3.33}$$

3.7. Podsumowanie

Stosowanie systemów doradczych wymaga posiadania reguł pozwalających na wnioskowanie o stanie obiektu na podstawie obserwowanych skutków tego stanu. Podstawą określania takich reguł mogą być relacje diagnostyczne.

Relacje diagnostyczne można identyfikować na podstawie eksperymentów diagnostycznych, wiąże się to jednak z pewnymi niedogodnościami: podczas biernego i czynno–biernego eksperymentu diagnostycznego obserwuje się niewielki zakres zmian stanu obiektu. Niedogodności tej pozbawiony jest eksperyment czynny, którego istotą jest celowe wprowadzanie zmian stanu obiektu, uniemożliwia to jednak stosowanie tego rodzaju eksperymentu w przypadku maszyn krytycznych jakimi są turbogeneratory energetyczne. Czynne eksperymenty diagnostyczne na obiekcie rzeczywistym można zastąpić badaniami symulacyjnymi czyli eksperymentami czynnymi na modelu symulatora obiektu. Zaawansowane numeryczne modele symulatorów obiektów pozwalają na przyjęcie założenia, że wyniki badań symulacyjnych można przenieść na obiekt rzeczywisty. Stosowanie takich badań do identyfikacji relacji diagnostycznych jest trudne ze względu na konieczność posiadania dużej wiedzy dotyczącej parametrycznego doboru wartości cech stanu w sposób umożliwiający otrzymanie określonego symptomu.

Poszukiwaną relacją diagnostyczną może być odwrotny model diagnostyczny, czyli model diagnostyczny uzyskany w wyniku operacji odwracania numerycznego modelu symulatora obiektu. Zasadniczą niedogodnością utrudniającą stosowanie modeli odwrotnych jest brak możliwości analitycznego odwracania znanych i zweryfikowanych złożonych numerycznych modeli symulatorów obiektów. Metodą umożliwiającą wyznaczenie odwrotnych modeli diagnostycznych jest metoda identyfikacji takich modeli na podstawie przykładów opisujących cechy stanu diagnozowanego obiektu oraz odpowiadające im cechy sygnałów diagnostycznych. Przykłady takie pozyskiwane są w wyniku badań symulacyjnych.

Modele odwrotne mogą być identyfikowane jako globalne lub lokalne. W [43] pokazano zastosowanie modeli globalnych identyfikowanych z użyciem sztucznych sieci neuronowych. Mimo uzyskania odpowiedniej jakości modelu odwrotnego stwierdzono małą przydatność modeli globalnych ze względu na ich dużą złożoność (wyznaczanie i doskonalenie takich modeli jest niezwykle czasochłonne) oraz utrudnienia związane z próbą rozszerzenia zbioru uwzględnianych przykładów. Niedogodności tych pozbawione są lokalne modele interpolacyjne. Modele te nie zastępują danych lecz są rozpięte na przykładach zawartych w zbiorach danych wzorcowych, wyznaczane są w chwili zapotrzebowania na podstawie aktualnych danych. Głównym czynnikiem decydującym o jakości tak identyfikowanych modeli jest równomierne wypełnienie przestrzeni symptomów przykładami wzorcowymi.

Podstawową trudnością identyfikacji odwrotnych modeli diagnostycznych na podstawie przykładów uzyskanych z numerycznych modeli obiektów jest występowanie nadmiernej liczby możliwych do wyznaczenia cech sygnałów diagnostycznych. Uwzględnianie wszystkich takich cech znacznie zwiększa rozmiar zadania identyfikacji modeli, ma również niekorzystny wpływ na jakość tych modeli. Rozwiązaniem tego problemu jest wybór zbiorów cech *relewantnych* (cech istotnych ze względu na rozpatrywany cel), które będą podstawą identyfikacji modeli odwrotnych.

Rozdział 4

Problem badawczy

W rozdziale 3 przedstawiono różne metody identyfikacji relacji diagnostycznych oraz zwrócono uwagę na problemy związane z ich zastosowaniem.

Metoda identyfikacji relacji diagnostycznych na podstawie badań symulacyjnych polega na takim sterowaniu badaniami symulacyjnymi (poprzez dobór odpowiednich wartości cech stanu), które umożliwi otrzymanie określonego symptomu. Sterowanie to może być wykonywane „ręcznie", wymaga to jednak dużej wiedzy i doświadczenia. Lepszym sposobem byłoby zastosowanie do tego celu odpowiedniego algorytmu. Ponieważ zadanie sterowania badaniami symulacyjnymi można sprowadzić do zadania wyboru odpowiednich cech stanu można to zadanie rozpatrywać jako zadanie optymalizacji zbioru takich cech.

Poszukiwaną relacją diagnostyczną może być odwrotny model diagnostyczny. Modele takie można identyfikować na podstawie przykładów (opisujących zadane stany rozpatrywanego obiektu oraz odpowiadające tym stanom symptomy) jako lokalne (ze względu na zakres stosowania) modele interpolacyjne. W rozdziale 3.6.3. zwrócono uwagę, że głównym czynnikiem decydującym o jakości tak identyfikowanych modeli jest równomierne wypełnienie przestrzeni symptomów przykładami wzorcowymi. Zadanie optymalizacji tak identyfikowanych modeli odwrotnych można sprowadzić do zadania optymalizacji zbiorów przykładów wzorcowych.

Kolejną trudnością związaną z identyfikacją modeli odwrotnych na podstawie przykładów uzyskanych z numerycznych modeli obiektów jest występowanie nadmiernej liczby możliwych do wyznaczenia cech sygnałów diagnostycznych. Uwzględnianie wszystkich możliwych do wyznaczenia cech znacznie zwiększa rozmiar zadania identyfikacji modeli odwrotnych. Jednym ze sposobów ograniczenia rozmiarów zadania jest dekompozycja przestrzenna modelu globalnego na modele lokalne związane z zespołami, podzespołami czy elementami maszyny. Dekompozycję tą należy połączyć z uniezależnieniem od siebie modeli lokalnych. Można to osiągnąć poprzez wprowadzenie sygnałów dodatkowych, wówczas model odwrotny identyfikowany jest jako dwustopniowy model mieszany (na pierwszy i drugi stopień składają się modele lokalne).

Modele lokalne stanowiące pierwszy stopień modelu odwrotnego przekształcają cechy sygnałów diagnostycznych w cechy sygnałów dodatkowych. Modele te mogą być

identyfikowane jako klasyfikatory wskazujące przynależność przykładów do klas sygnałów dodatkowych na podstawie wartości cech sygnałów diagnostycznych. Zadanie to jest odmienne od typowych zadań klasyfikacji ponieważ o przynależności do klas w danej przestrzeni decyduje się na podstawie wartości cech w innej przestrzeni. Zadanie to można rozwiązać poprzez poszukiwanie takiego zbioru cech sygnałów diagnostycznych, który zapewni zgodne rozmieszczenie przestrzenne przykładów uczących w przestrzeni tych cech z rozmieszczeniem w przestrzeni cech sygnałów dodatkowych, a dokładniej — zapewni uporządkowanie odległości pomiędzy przykładami w przestrzeni tych cech takie jak w przestrzeni cech sygnałów dodatkowych.

Stosowanie przedstawionych metod identyfikacji relacji diagnostycznych wymaga opracowania (bądź wyszukania) odpowiednich algorytmów umożliwiających:

- optymalizację zbioru wartości cech stanu,
- optymalizację zbioru przykładów,
- optymalizację zbioru cech sygnałów diagnostycznych.

Takimi algorytmami są algorytmy ewolucyjne, stosowane jako uniwersalne algorytmy optymalizacji umożliwiające znalezienie rozwiązania optymalnego dla różnych zadań optymalizacji.

4.1. Cel pracy

Celem pracy jest zbadanie możliwości zastosowania algorytmów ewolucyjnych w identyfikacji odwrotnych modeli diagnostycznych na podstawie przykładów uzyskanych w wyniku badań symulacyjnych z użyciem numerycznych modeli obiektów.

4.2. Tezy pracy

Sformułowano następujące tezy rozprawy:

1. **Poszukiwanie optymalnego zbioru cech dla identyfikacji odwrotnych modeli diagnostycznych prowadzi do lepszych rezultatów niż poszukiwanie zbioru cech optymalnych.**
2. **Nadmierne zwiększanie liczności zbioru cech, na podstawie których zostanie wyznaczony odwrotny model diagnostyczny, może prowadzić do pogorszenia jakości wyznaczonego modelu.**

Teza pierwsza ma na celu zwrócenie uwagi na różnicę występującą między wynikami działań, w których podczas dokonywania wyboru optymalnego zbioru cech uwzględnia się równocześnie wszystkie wybierane cechy oraz wynikami działań, polegających na zastosowaniu powszechnie znanych algorytmów przeszukiwania (w głąb, wszerz, z nawrotami

itp.) gdzie w każdym kroku wybierana jest jedna cecha. Różnica polega na sposobie oceniania cech tworzących zbiór: w pierwszym przypadku cechy te oceniane są łącznie (co pozwala na uwzględnienie wszystkich związków występujących miedzy nimi), w drugim — każda cecha oceniana jest oddzielnie.

Intuicyjnie uzasadnioną prawidłowością jest spostrzeżenie, że większa liczba poprawnie uwzględnianych cech obiektu powinna prowadzić do lepszej jakości wyznaczanego modelu. Z przeprowadzonych badań wstępnych wynika, że jest to prawdziwe wyłącznie dla małej liczby cech. Począwszy od pewnej granicznej liczby cech, uwzględnianie dodatkowych cech prowadzi do pogorszenia jakości wyznaczanego modelu. Ponieważ nie jest znany sposób określania tej liczby granicznej, w sformułowaniu drugiej z tez użyto nieprecyzyjnego określenia „nadmierne zwiększanie".

4.3. Zakres pracy

W rozdziale 2. omówiono algorytmy ewolucyjne. Przedstawiono ogólną koncepcję działania takich algorytmów. Zwrócono uwagę na zalety i wady rzeczywistoliczbowej oraz binarnej reprezentacji chromosomów. Zestawiono sposoby prowadzenia operacji selekcji, krzyżowania, mutacji jak również sukcesji właściwe dla wymienionych wyżej sposobów reprezentacji. Wskazano możliwości doboru innych parametrów algorytmu ewolucyjnego.

W rozdziale 3. przedstawiono różne koncepcje procesu diagnozowania obiektu. Wskazano na konieczność opracowania skutecznych algorytmów identyfikacji modeli diagnostycznych na potrzeby systemów doradczych. Omówiono zagadnienia związane z pozyskiwaniem relacji diagnostycznych. Wskazano na ich podstawowe znaczenie w diagnostyce maszyn. Przedstawiono problematykę modeli diagnostycznych oraz koncepcję modeli odwrotnych. Wskazano na dekompozycję modelu obiektu jako sposób ograniczenia rozmiarów zadania umożliwiający skuteczną identyfikację modeli odwrotnych. Omówiono wybrane algorytmy pozwalające na identyfikację i stosowanie takich modeli.

W rozdziale 4., na podstawie przeglądu stanu aktualnej wiedzy, sformułowano problem badawczy oraz tezy pracy.

Rozdział 5. zawiera opis zaproponowanej metody identyfikacji pierwszego stopnia dwustopniowego modelu diagnostycznego. Przedstawiono koncepcję odtwarzania klas stanów na podstawie wartości relewantnych cech sygnałów diagnostycznych. Pokazano zaproponowany sposób ewolucyjnego wyboru cech relewantnych. Omówiono sposób reprezentacji osobników przekształcanych przez algorytm ewolucyjny, operatory genetyczne oraz specjalnie opracowane warianty funkcji przystosowania.

Przykład zastosowania opracowanej metody opisano w rozdziale 6. W rozdziale tym przedstawiono obiekt badań oraz sposób jego dekompozycji poprzez wybrane sygnały dodatkowe. Badania przeprowadzono na podstawie przykładów uzyskanych w wyniku obliczeń symulacyjnych z użyciem opracowanego w IMP PAN w Gdańsku zaawansowanego modelu turbozespołu energetycznego dużej mocy.

Wyniki przeprowadzonych badań podsumowano w ostatnim rozdziale pracy. Przedstawiono możliwości doskonalenia opracowanej metody.

Rozdział 5

Zastosowanie algorytmów ewolucyjnych w identyfikacji modeli odwrotnych

5.1. Sterowanie badaniami symulacyjnymi z użyciem algorytmu ewolucyjnego

Identyfikacja relacji diagnostycznych na podstawie badań symulacyjnych wymaga takiego sterowania badaniami symulacyjnymi (poprzez dobór odpowiednich wartości cech stanu), które umożliwi otrzymanie określonych symptomów. Zastosowanie algorytmu ewolucyjnego do takiego sterowania badaniami symulacyjnymi pokazano w [111].

Zastosowano algorytm ewolucyjny, którego schemat pokazano na rys. 2.1. Algorytm przetwarza populację osobników. Genotyp osobnika tworzy pojedynczy chromosom składający się z genów, w których zapisane są wartości rozpatrywanych cech stanu. Chromosom ma postać macierzy jednokolumnowej, nie zastosowano kodowania — geny przyjmują wartości rzeczywiste cech stanu. Fenotypem osobnika jest zbiór wartości rozpatrywanych cech sygnałów diagnostycznych (symptomów) uzyskanych w wyniku badań symulacyjnych dla zapisanych w genotypie wartości cech stanu. Zastosowano operatory genetyczne opracowane dla przyjętego sposobu zapisu genotypu osobników.

Celem obliczeń było rozpoznanie stanu dla zadanych symptomów czyli znalezienie takich wartości cech stanu, które umożliwią uzyskanie w wyniku symulacji określonych symptomów. Mając to na uwadze, funkcję oceniającą określono jako funkcję podobieństwa tych symptomów i symptomów zapisanych w fenotypie osobnika. Prowadząc obliczenia dąży się do maksymalizacji funkcji oceniającej.

Przedstawione w artykule [111] wyniki obliczeń wskazują na wysoką przydatność algorytmów ewolucyjnych do sterowania badaniami symulacyjnymi — tak prowadzone badania umożliwiają znalezienie zbioru poszukiwanych wartości cech stanu.

Mając na uwadze możliwość zastosowania przedstawionej metody do pozyskiwania relacji służących do określania reguł na potrzeby systemów doradczych należy wskazać

jej następujące niedogodności:

- pozyskiwane są relacje typu (*wartości cech stanu* $\rightarrow$ *wartości cech sygnałów diagnostycznych*, określanie na ich podstawie reguł typu: **jeżeli** *symptom* **to** *klasa stanu* wymaga odpowiedniej wiedzy;

- długi czas obliczeń utrudnia stosowanie tej metody w systemach doradczych o ograniczonym i gwarantowanym czasie działania.

5.2. Ewolucyjna optymalizacja zbioru przykładów wzorcowych

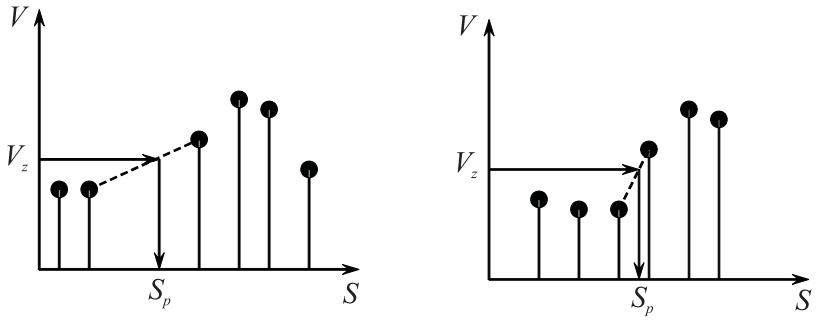
Długi czas obliczeń związanych z identyfikacją relacji diagnostycznych na podstawie badań symulacyjnych wynika z konieczności przeprowadzenia dużej liczby obliczeń symulacyjnych. Podczas obliczeń ewolucyjnych, w wyniku użycia operatorów genetycznych powstaje genotyp nowego osobnika (nowy zestaw wartości cech stanu), dla którego konieczne jest uzyskanie fenotypu (uzyskanie wartości cech sygnałów diagnostycznych w wyniku obliczeń symulacyjnych). Para (wartości cech stanu, odpowiadające im wartości cech sygnałów diagnostycznych) stanowi *przykład* danych symulacyjnych. Gromadzenie takich przykładów mogłoby skrócić czas obliczeń związanych z identyfikacją relacji diagnostycznych. Działanie takie jednak jest niecelowe ponieważ:

- stosowanie dokładnych wartości cech stanu i cech sygnałów diagnostycznych prowadzi do nieograniczonego zbioru przykładów,

- prawdopodobieństwo utworzenia genotypu, w wyniku operacji genetycznych, o wartosciach cech stanu takich jak w istniejącym już przykładzie jest bliskie zeru.

Metodą umożliwiającą identyfikację relacji diagnostycznych na podstawie przykładów jest metoda identyfikacji modeli odwrotnych jako lokalnych modeli interpolacyjnych. W metodzie tej zakłada się, że na podstawie pewnego zbioru przykładów (nazywanego zbiorem przykładów wzorcowych; zbiór ten jest podzbiorem ogólnego zbioru przykladów) w wyniku zastosowania algorytmu interpolacji można wyznaczyć poszukiwane wartości cech stanu S_p, którym odpowiadają zadane wartości cech sygnałów diagnostycznych V_z. Bezpośredni wpływ na jakość tak identyfikowanych modeli ma nie tylko dobór odpowiednich parametrów algorytmu interpolacji ale również odpowiednie wypełnienie przestrzeni symptomów przykładami wzorcowymi (rys. 5.1).

Zastosowanie algorytmów ewolucyjnych do optymalizacji zbioru przykładów wzorcowych zaproponowano w [38]. Zaproponowaną metodę obliczeń rozwinięto w [114]. W przedstawionej metodzie algorytm ewolucyjny przetwarza populację osobników zgodnie ze schematem pokazanym na rys. 2.1. Genotypem osobnika jest zbiór przykładów wzorcowych, fenotypem natomiast są wartości cech stanu wyznaczone na podstawie



Rys. 5.1. Wpływ rozmieszczenia przykładów wzorcowych na jakość modelu (na podstawie [76])

genotypu dla zadanych wartości cech sygnałów diagnostycznych. W wymienionych artykułach przedstawiono opracowane, dla tak przyjętej reprezentacji osobników, operatory genetyczne.

Obliczenia rozpoczynają się przy pustym ogólnym zbiorze przykładów, zbiór ten jest zapełniany przykładami wyznaczanymi w czasie obliczeń. W chwili zapotrzebowania na nowy przykład (w wyniku zadziałania operatorów genetycznych) w pierwszej kolejności sprawdzany jest zbiór ogólny, jeżeli poszukiwany przykład zostanie znaleziony nie ma potrzeby wykonywania obliczeń symulacyjnych. Nie wymaga się znalezienia dokładnie takiego przykładu jaki jest poszukiwany, wystarczy że znaleziony przykład leży w pewnym otoczeniu poszukiwanego przykładu. Taki sposób prowadzenia obliczeń powoduje, że po osiągnięciu pewnego stopnia zapełnienia przykładami zbioru ogólnego nie ma konieczności wyznaczania nowych przykładów, a optymalizacja zbioru przykładów wzorcowych sprowadza się do wybrania odpowiednich przykładów ze zbioru ogólnego. Prowadzi to do znacznego skrócenia czasu obliczeń co potwierdzają przedstawione wyniki badań.

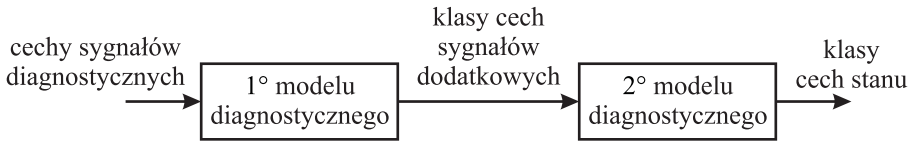
5.3. Ewolucyjna optymalizacja zbioru uwzględnianych cech

Opisane metody identyfikacji relacji diagnostycznych nie są metodami heurystycznymi [96] ponieważ nie umożliwiają uwzględnienia wiedzy o obiekcie badań. Wiedza ta jest zwykle fragmentaryczna co utrudnia stosowanie globalnych modeli heurystycznych. Model może być identyfikowany jako wielostopniowy gdzie jeden ze stopni jest modelem heurystycznym odpowiednio do posiadanej wiedzy o obiekcie. Najprostszym modelem wielostopniowym jest model dwustopniowy. Przekształcenie modelu jednostopniowego w model dwustopniowy jest możliwe dzięki wprowadzeniu tzw. *sygnałów dodatkowych* (rys. 3.14). Dobór odpowiednich sygnałów dodatkowych umożliwia również dekompozycję modelu globalnego na niezależne modele lokalne. Takie postępowanie znacznie zmniejsza rozmiary zadania identyfikacji modelu.

Pierwszy stopień modelu dwustopniowego przekształcający cechy sygnałów diagnostycznych w cechy sygnałów dodatkowych, może być traktowany jako pewien moduł

przetwarzania danych. Drugi stopień wiąże cechy sygnałów dodatkowych z cechami stanu. Sygnały dodatkowe powinny być tak dobrane by ich powiązanie z cechami stanu wynikało z wiedzy o obiekcie, co pozwoli na to aby drugi stopień modelu mógł być identyfikowany jako model heurystyczny.

Jak pokazano na rys. 3.13 nie można zakładać, że dokładny model odwrotny zawsze istnieje. Wynikających stąd trudności można uniknąć przez ograniczenie dokładności poszukiwanego modelu odwrotnego, zakładając, iż będzie wyznaczał on klasy wartości parametrów, a nie dokładne wartości tych parametrów (rys. 3.12). Tak identyfikowany model odwrotny można przedstawić jak na rys. 5.2.



Rys. 5.2. Dwustopniowy model diagnostyczny

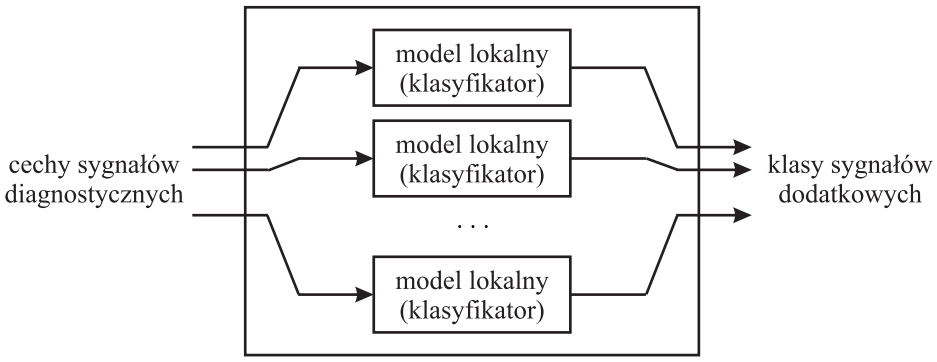
Zaproponowana metoda identyfikacji dwustopniowego modelu diagnostycznego składa się z następujących działań:

- dekompozycji globalnego modelu obiektu odpowiednio do wprowadzonych sygnałów dodatkowych,
- identyfikacji modeli lokalnych stanowiących pierwszy stopień (1°) modelu diagnostycznego; w zadaniu tym można wyróżnić podzadania:
  - wybór cech relewantnych,
  - wyznaczenie podziału w przestrzeni wartości cech sygnałów dodatkowych,
  - identyfikacja klasyfikatorów,
  - ocena jakości otrzymanych modeli,
- identyfikacji drugiego stopnia (2°) modelu diagnostycznego; zadanie to było przedmiotem badań prowadzonych przez J. Wojtusika i opisanych w [117].

5.3.1. Identyfikacja modeli lokalnych

Pierwszy stopień modelu diagnostycznego (stanowiące go modele lokalne) przekształca cechy sygnałów diagnostycznych w klasy cech sygnałów dodatkowych (rys. 3.14). Założono, że modele lokalne będą identyfikowane jako klasyfikatory, które na podstawie wartości cech sygnałów diagnostycznych rozpatrywanego przykładu określą stopnie jego przynależności do predefiniowanych klas cech sygnałów dodatkowych (rys. 5.3).

Jest to zadanie odmienne od typowego zadania klasyfikacji gdyż na podstawie wartości cech jednej przestrzeni (przestrzeni cech sygnałów diagnostycznych) określa się przynależność do klas zdefiniowanych w innej przestrzeni (przestrzeni cech sygnałów dodatkowych).

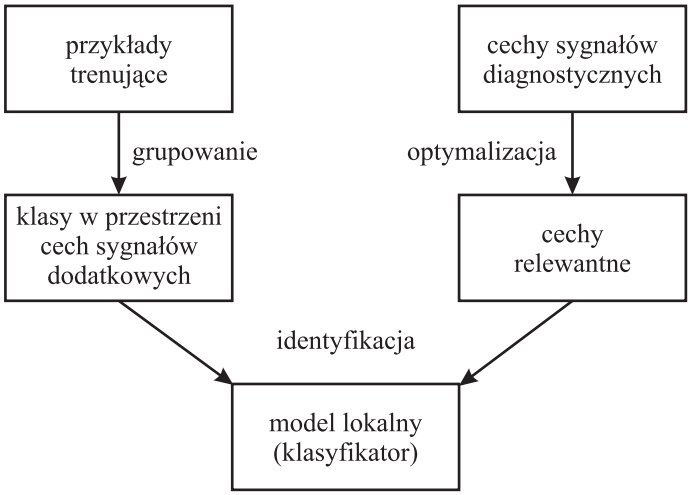


Rys. 5.3. 1° modelu diagnostycznego

Mając to na uwadze przyjęto, że optymalizacja tak określonych modeli lokalnych zostanie przeprowadzona poprzez optymalizację zbioru uwzględnianych cech sygnałów diagnostycznych. Założono, że wybrany optymalny zbiór cech (zbiór cech relewantnych), zapewni zgodny rozkład przykładów w przestrzeni cech sygnałów diagnostycznych oraz w przestrzeni cech sygnałów dodatkowych. Tak prowadzona optymalizacja nie wymaga przeprowadzenia wstępnego podziału na klasy.

Niezależnie od wyboru cech relewantnych należy wyznaczyć klasy w przestrzeni cech sygnałów dodatkowych. Wyznaczone klasy wraz z cechami relewantnymi stanowią podstawę do identyfikacji klasyfikatora będącego modelem lokalnym.

Istotę metody identyfikacji modeli lokalnych przedstawiono na rys. 5.4.



Rys. 5.4. Schemat metody identyfikacji modeli lokalnych

5.3.2. Wybór cech relewantnych

Najprostszym sposobem wyboru optymalnego podzbioru cech (zbioru cech relewantnych) ze zbioru wszystkich rozpatrywanych cech jest przeszukiwanie wyczerpujące. Metoda ta gwarantuje znalezienie najlepszego rozwiązania jednak czas obliczeń byłby nieakceptowalnie długi.

Metodą heurystycznego przeszukiwania dającą bardzo dobre wyniki w akceptowalnym czasie jest algorytm ewolucyjny. Mając to na uwadze zdecydowano o użyciu takiego algorytmu do wyboru zbioru cech relewantnych. Podstawową zaletą takiego podejścia jest jednoczesne poszukiwanie całego zbioru cech relewantnych, co zgodnie z przyjętą tezą powinno prowadzić do lepszych wyników niż wybór każdej z cech z osobna.

Dodatkowo, w celu porównania otrzymanych wyników oraz sprawdzenia słuszności przyjętej tezy, zdecydowano przeprowadzić poszukiwania zbioru cech relewantnych metodą wyboru z powracaniem oraz poprzez wybór zadanej liczby najlepszych cech ze zbioru wszystkich rozpatrywanych cech.

Zmniejszenia liczby rozpatrywanych cech można również osiągnąć stosując metody skalowania wielowymiarowego. Istotą tych metod jest przekształcenie zbioru wszystkich rozpatrywanych cech w zbiór nowych cech o mniejszej liczności tak by zachować relacje odległości elementów (przykładów) w obu przestrzeniach. Nowe cechy są kombinacjami liniowymi wszystkich rozpatrywanych cech. Zaletą tego podejścia jest automatyczne spełnienie kryterium oceny zbioru wybranych cech. Zaleta ta jest jednocześnie wadą ze względu na brak możliwości optymalizacji tego przekształcenia. Dodatkowo rozwiązanie to tylko częściowo spełnia przyjęte założenie o zmniejszeniu rozmiarów zadania identyfikacji pierwszego stopnia modelu diagnostycznego ponieważ otrzymanie nowych cech wymaga wyznaczenia wszystkich rozpatrywanych cech sygnałów diagnostycznych.

Ocena jakości zbioru cech relewantnych

Sposób oceny jakości zbioru cech relewantnych musi uwzględniać cel, dla którego zbiór ten jest poszukiwany. Celem tym jest otrzymanie rozmieszczenia przykładów uczących w przestrzeni tych cech podobnego do rozmieszczenia przykładów uczących w przestrzeni sygnałów dodatkowych — tak by na podstawie wartości wybranych cech można było wnioskować o klasie sygnałów dodatkowych.

Zaproponowano następujące sposoby oceny jakości zbiorów cech relewantnych:

1. na podstawie średniej liczby zgodnych najbliższych sąsiadów (w zadanym otoczeniu, czyli w zbiorze zadanej liczby ls najbliższych sąsiadów) w przestrzeni sygnałów dodatkowych oraz przestrzeni cech relewantnych:

$$oc_1 := \frac{1}{n} \sum_{i=1}^{n} \frac{lzs_i}{ls}, \tag{5.1}$$

   gdzie:
   lzs_i — liczba zgodnych najbliższych sąsiadów i–tego przykładu,
   n — liczba przykładów trenujących;
   na podstawie wstępnych badań przyjęto liczbę rozpatrywanych najbliższych sąsiadów $ls = 20$;

2. na podstawie zachowania proporcji odległości pomiędzy przykładami w obu przestrzeniach z użyciem miary *STRESS* (STandard REsidual Sum of Squares) metody

mapowania Sammon'a [66]:

$$oc_2 := \frac{1}{\sum_i \sum_{j<i} d^*_{ij}} \sum_i \sum_{j<i} \frac{(d^*_{ij} - d_{ij})^2}{d^*_{ij}}, \tag{5.2}$$

gdzie:
d^*_{ij} — odległość między przykładami i oraz j w przestrzeni sygnałów dodatkowych,
d_{ij} — odległość między przykładami i oraz j w przestrzeni cech relewantnych;

3. na podstawie wartości funkcji podobieństwa odległości *ccc* (Cophenetic Correlation Coeficient) [103, 104] między przykładami w przestrzeni cech relewantnych do odległości w dendrogramie wyznaczonym dla przykładów w przestrzeni cech sygnałów dodatkowych:

$$oc_3 := \frac{\sum_{i<j}(d_{ij} - \overline{d})(cd_{ij} - \overline{cd})}{\sqrt{\sum_{i<j}(d_{ij} - \overline{d})^2 \sum_{i<j}(cd_{ij} - \overline{cd})^2}}. \tag{5.3}$$

gdzie:
d_{ij} — odległość między przykładami i oraz j w przestrzeni cech relewantnych,
$\overline{d}$ — średnia odległość między przykładami w przestrzeni cech relewantnych,
cd_{ij} — odległość między przykładami i oraz j w dendrogramie utworzonym w przestrzeni sygnałów dodatkowych,
$\overline{d}$ — średnia odległość między przykładami w dendrogramie utworzonym w przestrzeni sygnałów dodatkowych.

Tak określone oceny przyjmują wartości z przedziału $[0; 1]$.

Algorytm ewolucyjny

Użycie algorytmu ewolucyjnego (algorytm 2.1) do rozwiązania rozpatrywanego zadania wymaga określenia: sposobu reprezentacji potencjalnych rozwiązań zadania (populacji przetwarzanej przez algorytm), sposobu selekcji rozwiązań rokujących, sposobu prowadzenia operacji genetycznych oraz warunku (warunków) zakończenia działania algorytmu.

Populacja W ogólnym zadaniu wyboru cech relewantnych nie jest również znana liczba wybieranych cech. Algorytm ewolucyjny umożliwia jednoczesną z wyborem cech optymalizację ich liczby. W celu realizacji takiego zadania określono populację w następujący sposób:

- populacja P przetwarzana przez algorytm ewolucyjny składa się z n osobników p,
- genotypem $p.g$ osobnika p jest chromosom będący ciągiem binarnym o długości l odpowiadającej liczbie rozpatrywanych cech punktowych; cechom tym nadane są numery porządkowe; znak '1' w chromosomie na pozycji a oznacza że cecha o numerze a jest cechą relawantną — w ten sposób algorytm ewolucyjny „decyduje" również o liczbie cech relewantnych.

- fenotypem $p.f$ osobnika p jest rozkład przykładów trenujacych w przestrzeni cech relewantnych (dla danego osobnika),
- populacja początkowa generowana jest losowo.

Liczba cech relewantnych l_r może być parametrem algorytmu ewolucyjnego. Należy zbadać jej wpływ na sprawność opracowanych klasyfikatorów. W tym celu przyjęto następującą strukturę genotypu osobnika: genotypem $p.g$ osobnika p jest chromosom zapisany w postaci macierzy jednowierszowej zawierającej l_r liczb z zakresu $[1, l_w]$ będących numerami cech wybranych jako relewantne (l_w jest liczbą rozpatrywanych cech punktowych). Populacja początkowa jest generowana losowo tzn. dla każdego osobnika generuje się losowo jednowierszową macierz numerów cech, która stanowi jego genotyp (numery cech wybierane są bez powtórzeń). Kolejność występowania numerów cech w chromosomie nie jest nośnikiem żadnej informacji.

Założono, że liczność populacji (zadana liczbą n) nie ulega zmianie w trakcie działania algorytmu ewolucyjnego.

Funkcja przystosowania Jako funkcje przystosowania przyjęto opisane w niniejszym rozdziale zależności (5.1), (5.2) oraz (5.3). Tak określone funkcje przystosowania przyjmują wartości z przedziału $[0; 1]$, gdzie 0 oznacza brak przystosowania a 1 maksymalne przystosowanie.

Selekcja i sukcesja Przyjęto, że stosowana będzie selekcja proporcjonalna (rozdział 2.5.), sukcesja natomiast będzie prowadzona według modelu elitarnego (rozdział 2.5.).

Operacje genetyczne Na podstawie przeprowadzonych badań ([112],[113]) przyjęto że krzyżowanie będzie prowadzone jako dwupunktowe (rozdz. 2.6.).

W pierwszej wersji algorytmu ewolucyjnego użyto operatora mutacji równomiernej (rozdz. 2.6.). W wersji drugiej, zastosowany operator mutacji jest modyfikacją operatora mutacji równomiernej dostosowującą go do przyjętego sposobu zapisu chromosomów: wartość przeznaczonego do mutacji genu jest zmieniana na losowo wybraną z przedziału $[1; l_w]$, gdzie l_w jest liczbą wszystkich rozpatrywanych cech punktowych. Prawdopodobieństwo mutacji p_m uzależniono od „utykania" algorytmu w maksimach lokalnych: jeżeli w kolejnym kroku algorytmu nie została osiągnięta poprawa wartości przystosowania najlepszego osobnika prawdopodobieństwo mutacji zostaje zwiększone o $0,01$ ale tylko do momentu osiągnięcia maksymalnej wartości ustalonej na poziomie $0,6$; jeżeli poprawa wartości przystosowania jest obserwowana wartość prawdopodobieństwa mutacji wraca do poziomu początkowego — $p_m = 0,1$.

Wybrane dla drugiej wersji algorytmu operatory genetyczne oraz przyjęty sposób zapisu chromosomów nie gwarantują, że powstałe w wyniku użycia tych operatorów chromosomy będą rozwiązaniami dopuszczalnymi. Niepoprawność chromosomu polega na powtórzeniu się wartości genów co oznacza wielokrotny wybór cechy do zbioru cech relewantnych. Aby zapobiec temu zjawisku opracowano operator naprawy uruchamiany po

operatorach krzyżowania i mutacji. Jego działanie polega na zastąpieniu powtarzających się wartości genów przez inne losowo wybrane z przedziału $[1; l_w]$.

Parametry algorytmu ewolucyjnego Liczbę osobników populacji równą 30 przyjęto kierując się zaleceniami podanymi w literaturze [1] oraz badaniami własnymi [38, 110, 112, 113, 114].

Na podstawie wyników wstępnych badań ([39, 41]) przyjęto dwa warunki zakończenia działania algorytmu ewolucyjnego:

- na podstawie kryterium minimalnej szybkości poprawy: jeżeli nie uda się otrzymać lepszego rozwiązania w zadanej liczbie 250 kolejnych pokoleń,
- na podstawie kryterium maksymalnego kosztu: po osiągnięciu zadanej liczby 500 pokoleń.

Przyjęte kryteria mają na celu ograniczenie kosztu uzyskania rozwiązania. Wartości liczbowe zostały dobrane na podstawie analizy wyników badań wstępnych, z której wynika, że zwiększanie wartości granicznych przyjętych warunków zatrzymania algorytmu ewolucyjnego nie przynosi poprawy wyników.

Wybór z powracaniem

Wybór z powracaniem jest metodą umożliwiającą wybór zadanej liczby cech relewantnych. Metoda ta jest połączeniem metod regresji krokowej postępującej oraz wstecznej [46, 49, 55] stosowanych naprzemiennie: dwie cechy są dołączane do zbioru cech relewantnych metodą regresji krokowej postępującej, następnie jedna cecha z tego zbioru jest usuwana wg metody regresji krokowej wstecznej; działania te są powtarzane aż do osiągnięcia zadanej liczności zbioru cech relewantnych.

Dodanie cechy do zbioru cech relewantnych przebiega następująco:

- kolejno, każdą z wszystkich rozpatrywanych cech niewybraną do zbioru cech relewantnych, dołącza się tymczasowo do zbioru cech relewantnych w celu wyznaczenia jakości tak powiększonego zbioru,
- cechą relewantną zostaje ta cecha, której dołączenie do zbioru cech relewantnych dało najlepszą ocenę jakości tego zbioru.

Usuwanie cechy ze zbioru cech relewantnych również przebiega w dwóch krokach:

- kolejno, każda z cech relewantnych jest usuwana tymczasowo ze zbioru cech relewantnych w celu wyznaczenia jakości pomniejszonego zbioru,
- usunięta zostaje ta cecha, której brak w zbiorze cech relewantnych spowodował najmniejszy spadek jakości tego zbioru.

Wybór najlepszych cech

Wybór zadanej liczby l_r najlepszych cech ze zbioru wszystkich rozpatrywanych cech przebiega następująco:

- każdą cechę z osobna traktuje się jako jednoelementowy zbiór cech relewantnych i wyznacza się jakość tak powstałych zbiorów,
- szereguje się cechy wg jakości w porządku malejącym,
- l_r początkowych cech tworzy poszukiwany zbiór cech relewantnych.

5.3.3. Podział w przestrzeni sygnałów dodatkowych

Wyznaczenie klas w przestrzeni cech sygnałów dodatkowych jest kolejnym krokiem w procesie identyfikacji modeli lokalnych. Klasy te mogą być wyznaczone na podstawie wyników grupowania przykładów uzyskanych z zastosowaniem znanych algorytmów grupowania bądź też z uwzględniem wiedzy o obiekcie badań.

Przyjęto, że klasy w przestrzeni cech sygnałów dodatkowych, zostaną wyznaczone na podstawie podziału przykładów trenujących. Wstępnie założono, że podział ten zostanie uzyskany w wyniku grupowania [39] z zastosowaniem takich metod jak metoda hierarchicznego [52] łączenia czy k–najbliższych sąsiadów [119].

5.3.4. Klasyfikatory

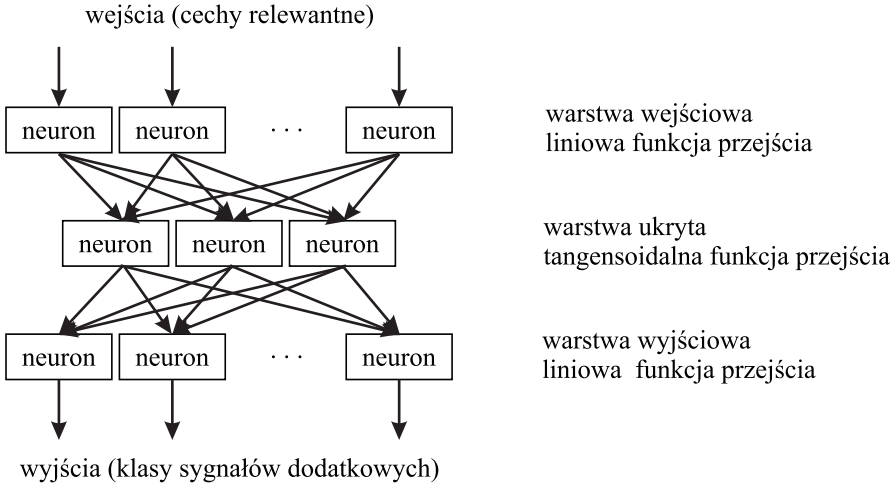
Zadaniem rozpatrywanego klasyfikatora jest klasyfikacja przykładów, na podstawie wybranych cech sygnałów, do klas sygnałów dodatkowych. Zdecydowano zastosować następujące klasyfikatory:

1. sztuczna sieć neuronowa,
2. algorytm SVM,
3. prosty klasyfikator.

Sztuczna sieć neuronowa

Budowę stosowanej sztucznej sieci neuronowej pokazano na rys. 5.5. Sieć składa się z trzech warstw:

- wejściowej o liniowej funkcji przejścia neuronów,
- ukrytej o tangensoidalnej funkcji przejścia neuronów,
- wyjściowej o liniowej funkcji przejścia neuronów.



Rys. 5.5. Budowa stosowanej sztucznej sieci neuronowej

Liczba neuronów warstwy wejściowej jest równa liczbie rozpatrywanych cech relewantnych. Warstwa ukryta składa się z trzech neuronów. Liczba neuronów warstwy wyjściowej jest równa liczbie przyjętych klas sygnałów dodatkowych.

Przyjęto, że dane trenujące są przygotowywane następująco:

- macierz wejść o wymiarach $m \times n$, m jest liczbą przykładów trenujących a n liczbą cech relewantnych,

- macierz wyjść o wymiarach $m \times k$, k jest liczbą klas sygnałów dodatkowych; cyfra 1 oznacza przynależność do danej klasy, 0 — brak przynależności.

Trenowanie sieci przebiega według algorytmu wstecznej propagacji błędów [107].

Budowę sztucznej sieci neuronowej oraz jej parametry dobrano na podstawie wstępnych badań.

Algorytm SVM

Standardowy algorytm SVM pozwala na budowę klasyfikatora binarnego, czyli umożliwiającego wskazanie przykładów należących do danej klasy oraz do niej nie należących. Stosowanie algorytmu SVM wymaga więc wyznaczenia klasyfikatora dla każdej klasy sygnałów dodatkowych.

Na podstawie wstępnych badań założono stosowanie algorytmu SVM z radialną funkcją jądra (3.20).

Prosty klasyfikator

Prosty klasyfikator tworzą reprezentanci klas. Reprezentantem klasy jest jeden z tworzących ją przykładów leżący najbliżej środka ciężkości klasy. Klasyfikacja prowadzona

jest na podstawie odległości klasyfikowanego przykładu od reprezentantów klas — przykład zostaje zaklasyfikowany do klasy, której reprezentant jest najmniej odległy od rozpatrywanego przykładu.

5.3.5. Ocena jakości otrzymanych modeli

Ocena jakości opracowanych modeli zostanie wyznaczona na podstawie uzyskanych dla nich wyników klasyfikacji względem wyników klasyfikacji w dziedzinie sygnałów dodatkowych. Wyznaczone zostaną następujące liczby:

- O_1 — względna liczba poprawnie zaklasyfikowanych przykładów,
- O_2 — względna liczba przykładów, które powinny zostać zaklasyfikowane do jednej z klas niesprawności, a zostały zaklasyfikowane do klasy: „brak niesprawności" (błąd typu nie wykrycie niesprawności),
- O_3 — względna liczba przykładów, które powinny zostać zaklasyfikowane do klasy: „brak niesprawności" a zostały zaklasyfikowane do jednej z klas niesprawności (błąd typu wykrycie nieistniejącej niesprawności),
- O_4 — względna liczba przykładów zaklasyfikowanych do jednej z klas niesprawności, które powinny być zaklasyfikowane do innej z tych klas (błąd polegający na wykryciu niewłaściwej niesprawności).

Rozdział 6

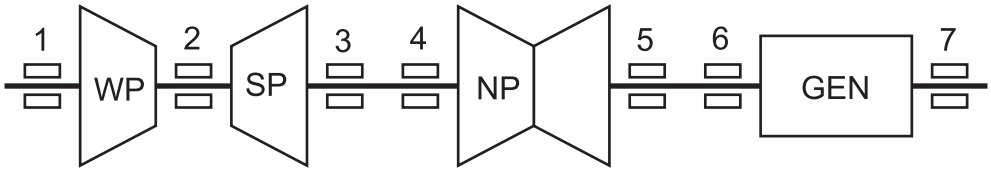
Przykład zastosowania opracowanej metody

W rozdziale 5.3. przedstawiono proponowaną przez autora metodę identyfikacji modeli lokalnych stanowiących pierwszy stopień dwustopniowego modelu odwrotnego. Zasadniczym elementem tej metody jest wybór cech relewantnych. Wyboru tego można dokonać używając m.in. zaproponowanego algorytmu ewolucyjnego. Stwierdzenie wysokiej sprawności klasyfikatorów oraz uzyskanie najlepszych wyników z użyciem algorytmu ewolucyjnego do wyboru ech relewantnych może być traktowane jako indukcyjny dowód postawionych tez oraz poprawności sformułowanych metod.

6.1. Obiekt badań

Badania przeprowadzono dla wybranego obiektu — turbogeneratora z turbiną 13K215, którego schemat pokazano na rys. 6.1. Na wybór takiego obiektu badań miały wpływ m.in. następujące przesłanki:

1. turbiny tego typu są powszechnie używane w krajowej energetyce, dzięki temu praca ta ma charakter utylitarny;
2. uzyskano dostęp do wyników badań symulacyjnych na opracowanym w IMP PAN w Gdańsku zaawansowanym modelu numerycznym takiego obiektu [74].



Rys. 6.1. Szkic poglądowy turbogeneratora z turbiną 13K215 [74]

Jako rozpatrywaną niesprawność przyjęto przemieszczenia podpór łożyskowych względem wymaganej postaci kinetostatycznej linii wału. Na podstawie analizy przyczyn występowania poważnych awarii maszyn wirnikowych można stwierdzić, że w ok.

80% dotyczą one wirnika i zespołu łożysk maszyny [16]. Jedną z głównych przyczyn tych awarii są właśnie przemieszczenia podpór łożyskowych. Nieodpowiednia postać kinetostatycznej linii wału powoduje zwiększenie sił promieniowych w układzie wirnik–łożyska co może prowadzić do znacznego skrócenia okresu międzyremontowego.

Postać kinetostatycznej linii wału była przedmiotem badań przedstawionych w [80]. W pracy tej przedstawiono metodę pozyskiwania relacji diagnostycznych bazującą na metodach uczenia maszynowego. Uzyskano wysoką sprawność klasyfikatorów przemieszczeń pojedynczych łożysk. Zaobserwowano znacznie niższą sprawność klasyfikatorów stosowanych podczas występowania przemieszczeń innych łożysk niż łożysko, którego położenia są klasyfikowane. Przykłady pozyskiwano na podstawie odpowiednio zaplanowanych badań symulacyjnych. Postępowanie takie pozytywnie wpływa na jakość pozyskiwanych relacji lecz znacznie wydłuża czas obliczeń co może w pewnych przypadkach uniemożliwić stosowanie przedstawionej metody. Problem związany z występowaniem nadmiernej liczby możliwych do wyznaczenia cech sygnałów diagnostycznych rozwiązano w [80] stosując trzyetapową redukcję liczby cech:

1. na podstawie statystycznej istotności cech,
2. na podstawie wartości współczynnika korelacji
3. na podstawie występowania zależności funkcyjnych.

Identyfikacja przemieszczeń podpór łożyskowych turbogeneratora była także przedmiotem badań opisanych w [117]. W pracy tej przedstawiono metodę identyfikacji wielomodeli stanowiących drugi stopień dwustopniowego modelu odwrotnego (pierwszym stopniem tego modelu odwrotnego były klasyfikatory opracowane w ramach niniejszej rozprawy). Weryfikację opracowanych wielomodeli przeprowadzono bazując na wynikach badań prowadzonych w ramach niniejszej rozprawy. Uzyskano wysoką sprawność opracowanych wielomodeli w wykrywaniu przemieszczeń nie tylko pojedynczych łożysk ale również jednoczesnego przemieszczenia dwóch łożysk.

6.1.1. Dane uczące

Dla wybranego obiektu badań uzyskano z IMP PAN zbiór danych symulacyjnych wyznaczonych z użyciem zaawansowanego modelu numerycznego turbiny i generatora. Ze względu na niewielką liczność uzyskanego zbioru danych, zbiór ten w całości został użyty jako zbiór danych uczących. Zbiór danych uczących został podzielony na dwa zbiory:

- zbiór danych trenujących zawierający 57 przykładów:
  - stan bazowy [99]: 1 przypadek,
  - przemieszczenia małe — zestaw 1 [99]: 4 kierunki * 7 łożysk = 28 przypadków,

  - przemieszczenia duże — zestaw 3 [99]: 4 kierunki * 7 łożysk = 28 przypadków,

- zbiór danych testujących zawierający 32 przykłady:

  - jednoczesne przemieszczenia łożyska 5 i 6 — zestaw 2 [99]: 4 kierunki * 2 łożyska * 4 kierunki = 32 przypadki.

Strukturę danych symulacyjnych przyjęto za [30]. Dla każdego przykładu uwzględniono 316 punktowych cech sygnałów wyznaczonych niezależnie dla każdego z siedmiu łożysk. Łącznie rozpatrywano 2212 cech punktowych.

6.2. Plan badań

Weryfikację opracowanej metody przeprowadzono w następujących krokach:

1. przeprowadzono dekompozycję modelu globalnego na modele lokalne wprowadzając wybrane sygnały dodatkowe,
2. wyznaczono klasy w przestrzeni sygnałów dodatkowych na podstawie podziału przykładów trenujących,
3. na podstawie danych trenujących poszukiwano zbiorów cech relewantnych,
4. dla wyszukanych zbiorów cech relewantnych przeprowadzono identyfikację modeli lokalnych (klasyfikatorów),
5. przeprowadzono klasyfikację danych trenujących w celu sprawdzenia dopasowania zidentyfikowanych modeli do danych trenujących,
6. przeprowadzono klasyfikację danych testujących w celu sprawdzenia skuteczności opracowanej metody.

Obliczenia związane z wyborem cech relewantnych oraz wyznaczeniem i weryfikacją klasyfikatorów przeprowadzono dla wszystkich wariantów opracowanej metody (patrz rozdz. 5.3.), dla których przyjęto trzycyfrowe oznaczenia:

- pierwsza cyfra oznacza metodę poszukiwania zbioru cech relewantnych:
  1. ewolucyjna optymalizacja liczby i rodzaju cech,
  2. ewolucyjna optymalizacja zbioru cech o zadanej liczności,
  3. wybieranie z powracaniem zadanej liczby cech,
  4. wybór zadanej liczby najlepszych cech;
- druga cyfra oznacza sposób oceny potencjalnych rozwiązań:

1. na podstawie zależności (5.1), w przestrzeni cech relewantnych użyto standaryzowanej miary Euklidesa, w przestrzeni reakcji podporowych — miary Mahalanobisa (decyzje te podjęto na podstawie wstępnych badań),
2. na podstawie zależności (5.2), w obu przestrzeniach zastosowano standaryzowaną miarę Euklidesa,
3. na podstawie zależności (5.3);

- trzecia cyfra oznacza rodzaj zastosowanego klasyfikatora:

1. sztuczna sieć neuronowa,
2. algorytm SVM,
3. prosty klasyfikator.

Dla każdego z wymienionych wariantów, w których liczba cech jest parametrem, przeprowadzono obliczenia dla następujących liczności zbioru cech relewantnych: 4, 5, 6, 7, 8, 9 oraz 10.

Obliczenia z użyciem algorytmu ewolucyjnego przeprowadzono dla każdego wariantu 20-krotnie, miało to na celu sprawdzenie stopnia powtarzalności wyników; dla pozostałych metod obliczenia każdego z wariantów przeprowadzono jednokrotnie — wyniki tych metod są w pełni powtarzalne.

6.2.1. Opracowane oprogramowanie

Zaplanowane badania przeprowadzono w środowisku MATLAB. W tym celu opracowano własne oprogramowanie w postaci funkcji tego pakietu realizujących poszczególne zadania:

- wyznaczenia podziału w przestrzeni sygnałów dodatkowych,
- klasyfikacji przykładów uczących do wyznaczonych klas z przedstawieniem wyników w formie rysunkowej,
- wybór zbioru cech relewantnych (algorytmy: ewolucyjny, wyboru z powracaniem, wyboru najlepszych cech)
- wyznaczenia klasyfikatorów oraz ich oceny.

Zadania identyfikacji klasyfikatorów z użyciem sztucznych sieci neuronowych oraz algorytmu SVM przeprowadzono stosując standardowe funkcje pakietu MATLAB.

6.3. Przebieg badań

6.3.1. Dekompozycja modelu, wybór sygnałów dodatkowych

Dla rozpatrywanego obiektu badań (turbina z generatorem) oraz wybranego rodzaju niesprawności tego obiektu (przemieszczenia podpór łożyskowych) zdecydowano się prze-

prowadzić dekompozycję modelu globalnego na modele lokalne związane z podporami łożyskowymi.

Przeprowadzenie dekompozycji modelu globalnego na modele lokalne jest celowe jedynie wtedy gdy połączone jest z rozprzęgnięciem modeli lokalnych. Rozprzęgnięcie to można uzyskać poprzez wprowadzenie odpowiednich sygnałów dodatkowych. Dla rozpatrywanego obiektu jako sygnały dodatkowe wybrano reakcje podporowe [41]. Wielkości te nie są obserwowane bezpośrednio w czasie eksploatacji obiektu. Zostały one wyznaczone w czasie badań symulacyjnych.

Na wybór takich sygnałów dodatkowych miały wpływ następujące przesłanki:

- przemieszczenia podpór wpływają na reakcje podporowe, które z kolei są przyczyną drgań tychże podpór, wobec tego należy oczekiwać, że możliwe będzie znalezienie 'przejścia' od drgań podpór (sygnałów diagnostycznych) poprzez reakcje podporowe (sygnały dodatkowe) do przemieszczeń podpór (stany),

- na wartości reakcji danej podpory wpływają nie tylko przemieszczenia rozpatrywanej podpory ale również przemieszczenia sąsiednich podpór — pozwala to uwzględnić wpływ przemieszczenia jednej z podpór na stan podpór sąsiednich,

- można założyć, że reakcje obserwowane dla danej podpory łożyskowej warunkują drgania tej i tylko tej podpory — to założenie pozwala na 'rozcięcie' modeli lokalnych, tzn. na rozpatrywanie tych modeli jako niezależnych.

6.3.2. Podział w przestrzeni sygnałów dodatkowych

Założono, że modele lokalne będą identyfikowane dla każdej podpory łożyskowej jako klasyfikatory pozwalające na wyznaczenie, na podstawie wybranych cech sygnałów, przynależności do predefiniowanych klas reakcji podporowych, które określono następująco:

- S — nominalna wartość reakcji podporowej (rys. 6.2 a),

- G — koniec wektora reakcji przesunięty w górę względem położenia nominalnego (rys. 6.2 b),

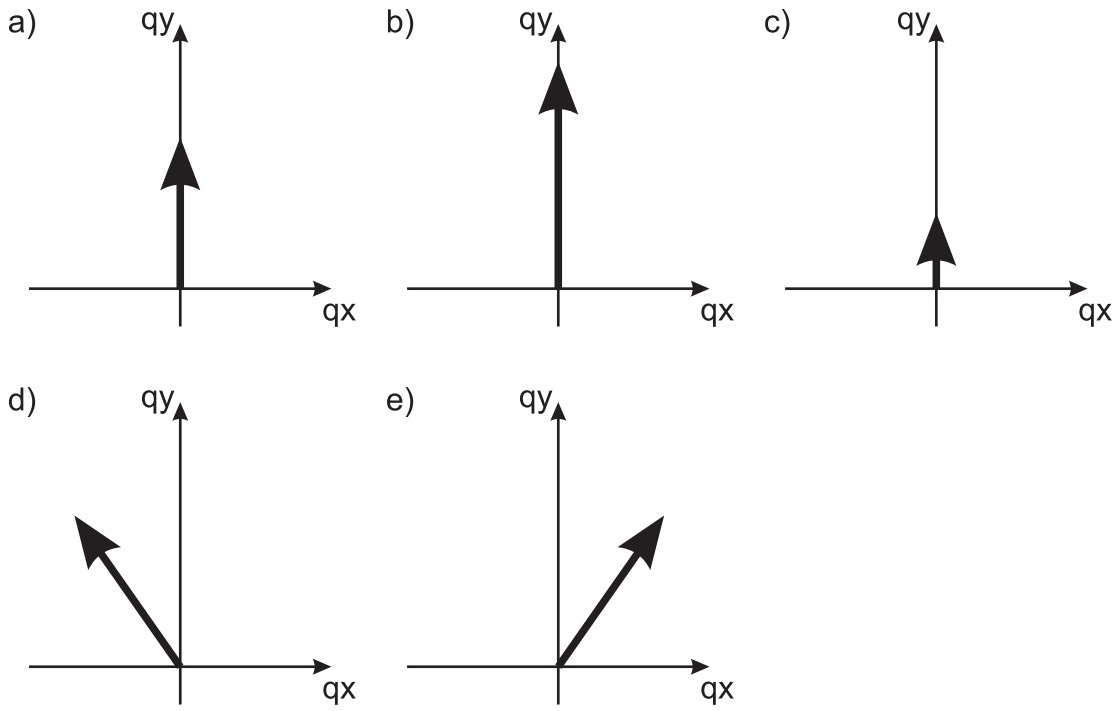
- D — koniec wektora reakcji przesunięty w dół (rys. 6.2 c),

- L — koniec wektora reakcji przesunięty w lewo (rys. 6.2 d),

- P — koniec wektora reakcji przesunięty w prawo (rys. 6.2 e).

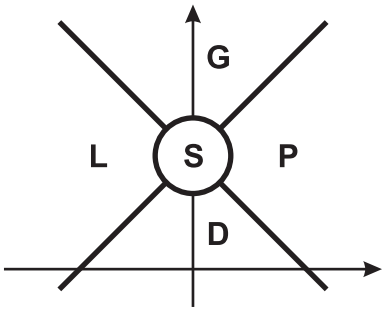
Klasy G, D, L oraz P oznaczają występowanie niesprawności natomiast klasa S — brak niesprawności.

Przyjęto, że klasy, na podstawie których zostanie opracowany klasyfikator, zostaną wyznaczone na podstawie podziału przykładów uczących w przestrzeni reakcji podporowych. Wstępnie założono, że podział ten zostanie uzyskany w wyniku grupowania



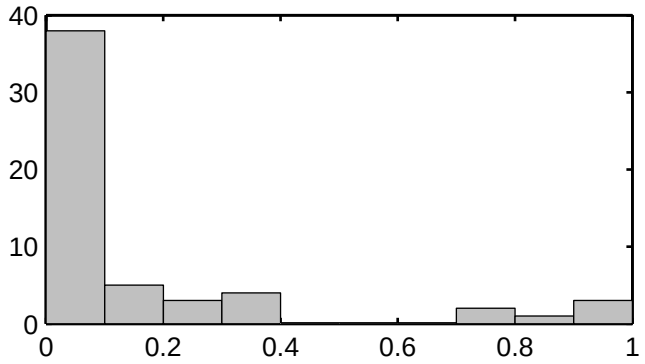
Rys. 6.2. Przykłady reakcji podporowych należących do przyjętych klas: a) klasa S, b) klasa G, c) klasa D, d) klasa L, e) klasa P

przykładów uczących [39]. Zastosowanie takich metod jak metoda hierarchicznego łączenia [52] czy k–najbliższych sąsiadów [119] nie przyniosło dobrych rezultatów. Otrzymane wyniki nie były zgodne z podziałem intuicyjnym. Grupowanie bez nadzoru stosuje się wtedy gdy chcemy odkryć pewne zależności występujące w danych. W tym przypadku nie ma takiej potrzeby. Znany jest sposób podziału, a głównym celem zastosowania jakiejś metody numerycznej jest dokonanie podziału w sposób zautomatyzowany — podział 'ręczny' byłby bardzo czasochłonny. Wobec tego zdecydowano się zastosować klasyfikację do klas zdefiniowanych przez określenie granic pomiędzy nimi (rys. 6.3): klasę S ogranicza okrąg, którego środkiem jest przykład bazowy, pozostałe klasy rozdzielone są prostymi prostopadłymi, przecinającymi się w punkcie określonym przez przykład bazowy, tworzącymi z osią poziomą kąt odpowiednio 45° i 135°.

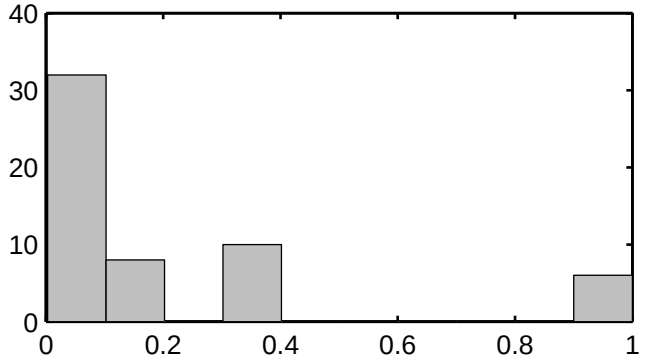


Rys. 6.3. Klasy w przestrzeni reakcji podporowych

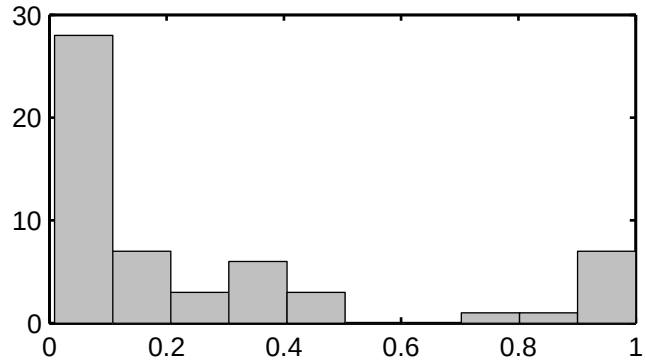
Taki sposób określenia klas przyjęto na podstawie wzrokowej analizy rozkładu przykładów trenujących w przestrzeni reakcji podporowych. Długość promienia okręgu ograniczającego klasę S określono na podstawie analizy histogramów względnych odległości przykładów trenujących od przykładu bazowego (rys. 6.4–6.10). Histogramy utworzono dzieląc zakres zmienności względnej odległości na 10 przedziałów. Największe zagęszczenie przykładów przypada na pierwszy przedział histogramów, na tej podstawie przyjęto długość promienia okręgu ograniczającego klasę S jako 10% maksymalnej odległości.



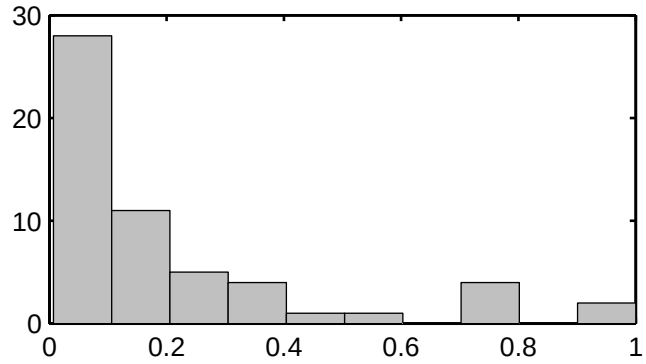
Rys. 6.4. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 1



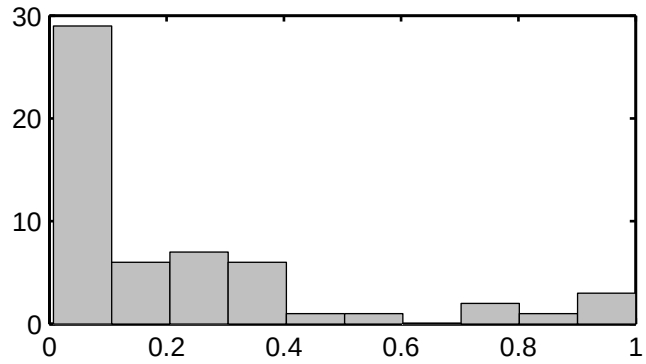
Rys. 6.5. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 2



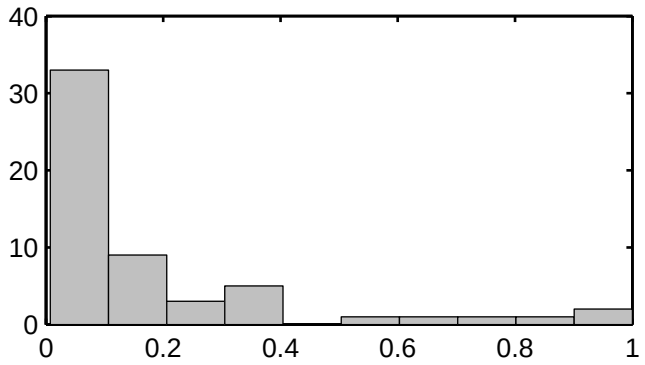
Rys. 6.6. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 3



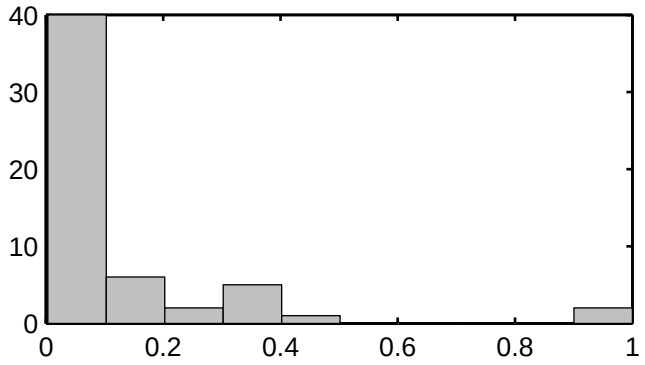
Rys. 6.7. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 4



Rys. 6.8. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 5



Rys. 6.9. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 6

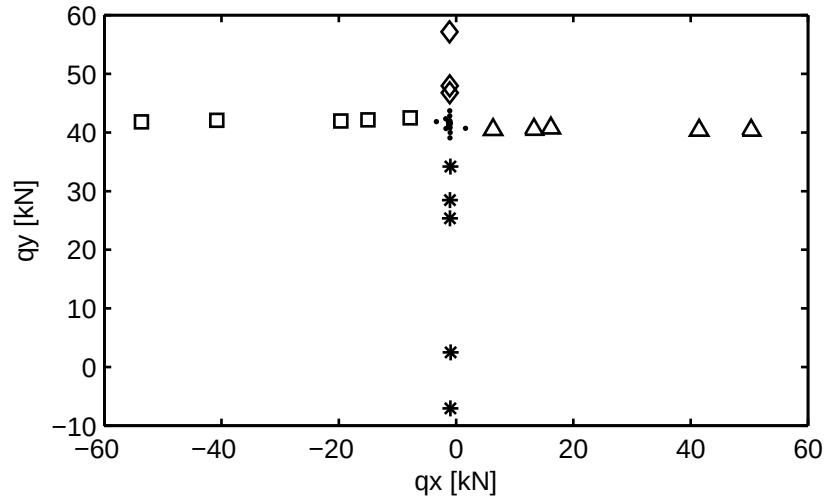


Rys. 6.10. Histogram względnych odległości przykładów trenujących od przykładu bazowego dla łożyska 7

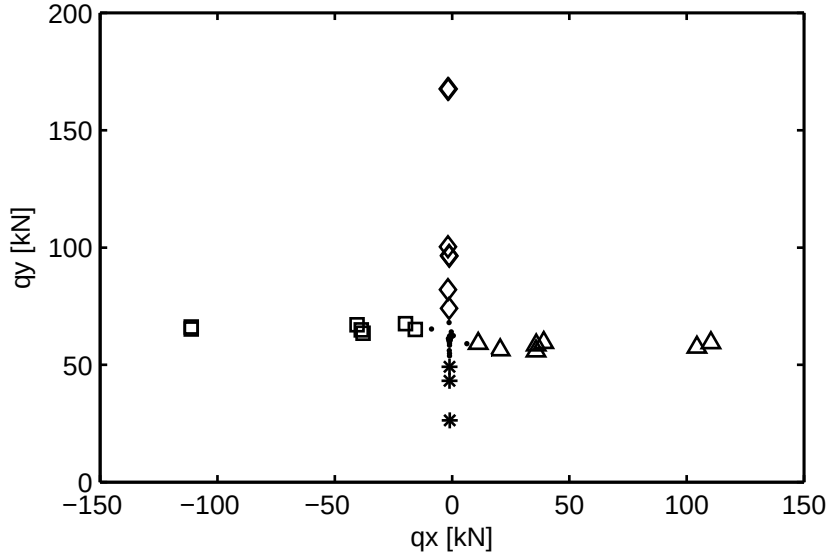
Na rys. 6.11–6.17 pokazano przynależność przykładów trenujących a na rys. 6.18–6.24 przykładów testujących do tak zdefiniowanych klas reakcji podporowych. Odpowiednimi symbolami zaznaczono przynależność przykładów trenujących i testujących do poszczególnych klas:

- $\Diamond$ — koniec wektora reakcji przesunięty w górę względem położenia nominalnego (klasa G),
- $*$ — koniec wektora reakcji przesunięty w dół (klasa D),
- $\Box$ — koniec wektora reakcji przesunięty w lewo (klasa L),
- $\triangle$ — koniec wektora reakcji przesunięty w prawo (klasa P),
- $\bullet$ — nominalna wartość reakcji podporowej (klasa S).

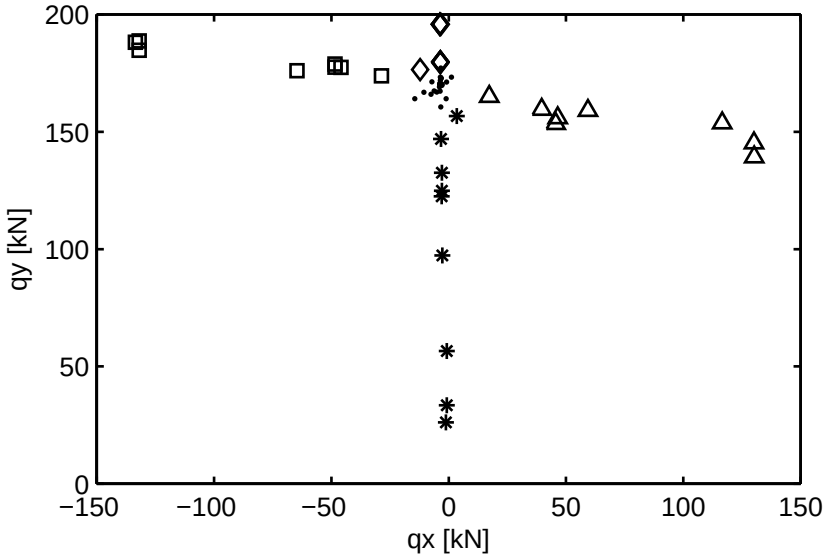
Na osi poziomej, przedstawionych na tych rysunkach wykresów, zaznaczono wartości składowej poziomej reakcji, zaś na osi pionowej wartości składowej pionowej reakcji dla danego łożyska turbogeneratora. Występowanie nominalnych wartości reakcji podporowych we wszystkich łożyskach turbogeneratora oznacza m.in. brak odstępstw od nominalnej linii kinetostatycznej wału, co oznacza poprawną (w tym zakresie) pracę maszyny. Odchylenie od nominalnych wartości reakcji podporowych oznacza wystąpienie niesprawności. Niesprawnością taką może być miedzy innymi przemieszczenie podpory łożyskowej (podpór łożyskowych), tego typu niesprawności rozpatrywane sa w pracy.



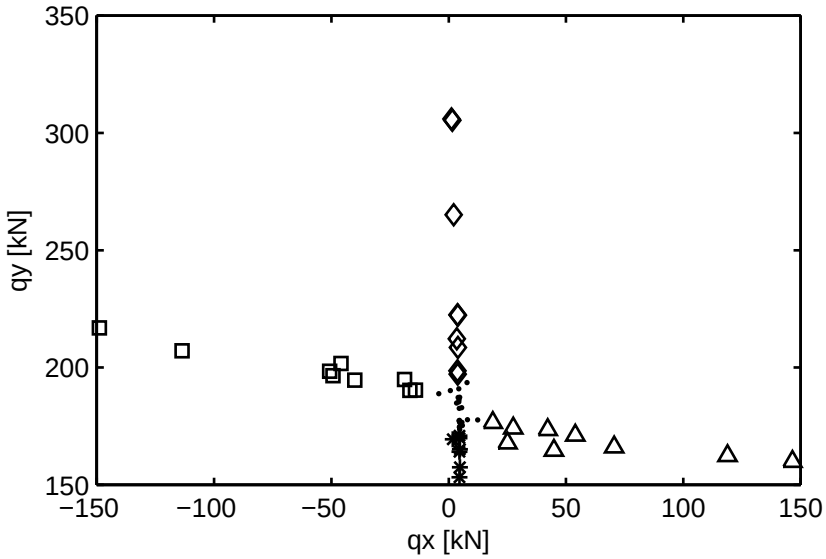
Rys. 6.11. Przykłady trenujące — klasy reakcji podporowych, łożysko 1



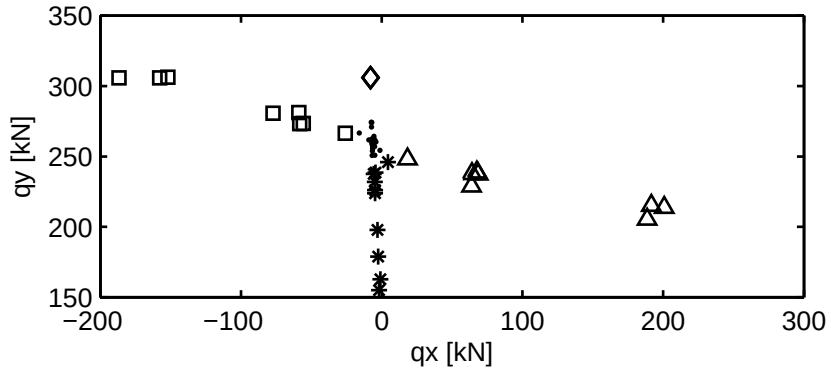
Rys. 6.12. Przykłady trenujące — klasy reakcji podporowych, łożysko 2



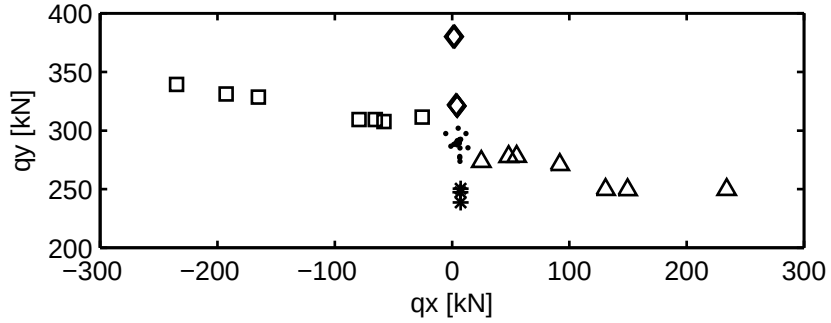
Rys. 6.13. Przykłady trenujące — klasy reakcji podporowych, łożysko 3



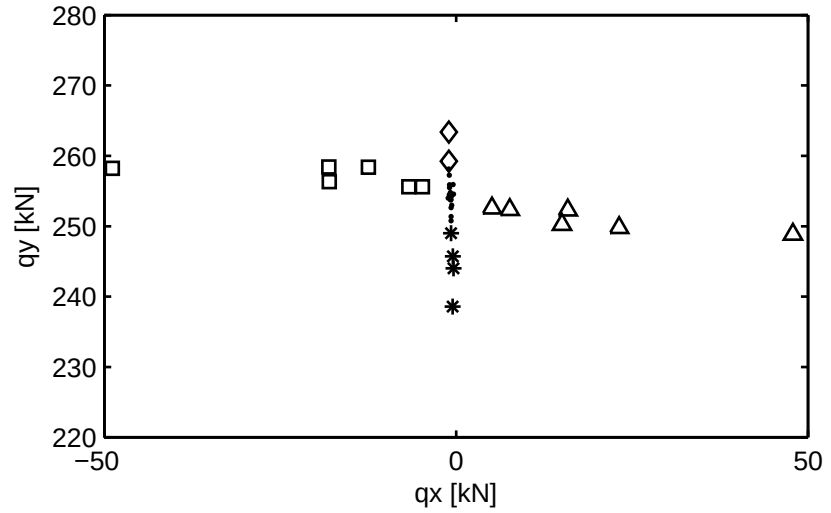
Rys. 6.14. Przykłady trenujące — klasy reakcji podporowych, łożysko 4



Rys. 6.15. Przykłady trenujące — klasy reakcji podporowych, łożysko 5

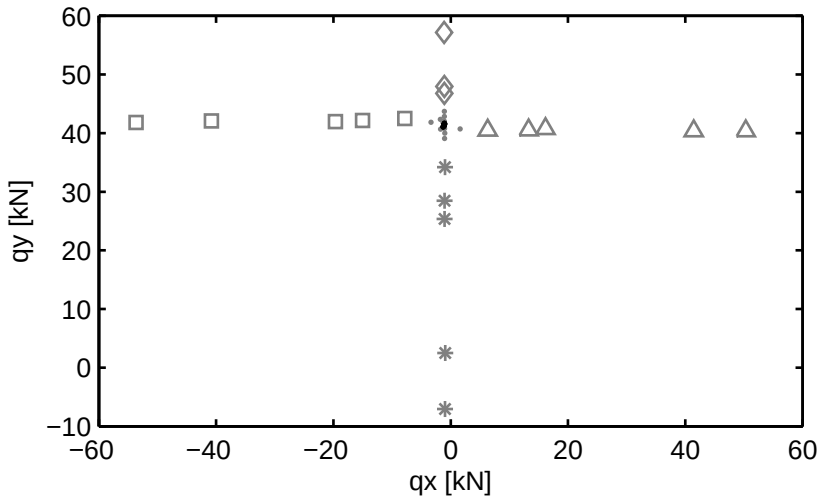


Rys. 6.16. Przykłady trenujące — klasy reakcji podporowych, łożysko 6

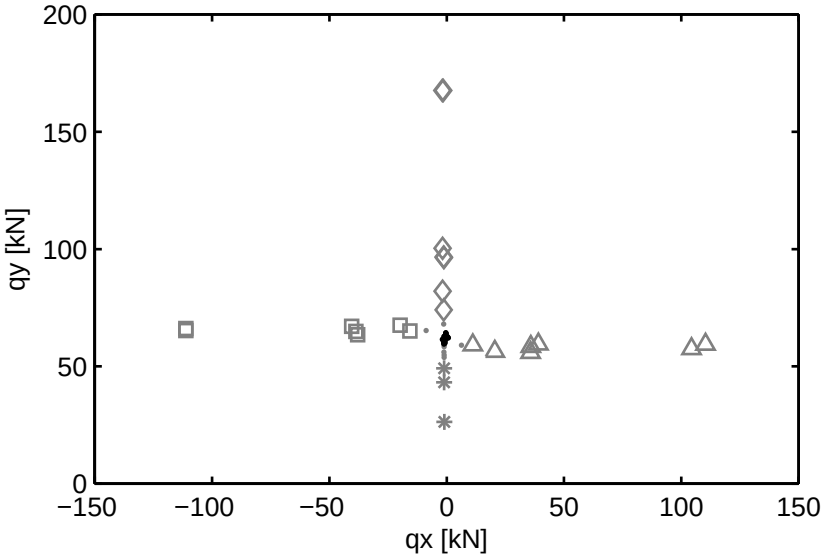


Rys. 6.17. Przykłady trenujące — klasy reakcji podporowych, łożysko 7

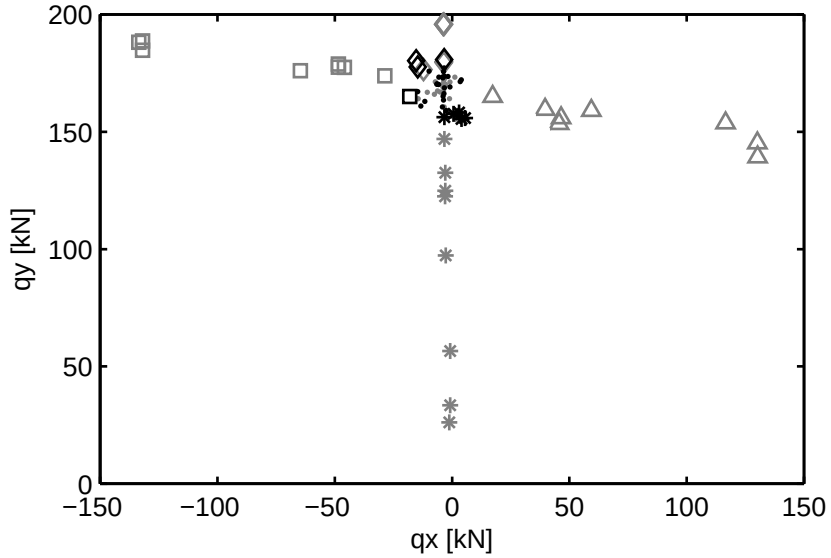
Na rysunkach 6.18–6.24 dotyczących danych testujących, w tle, szarym kolorem zaznaczono dane trenujące.



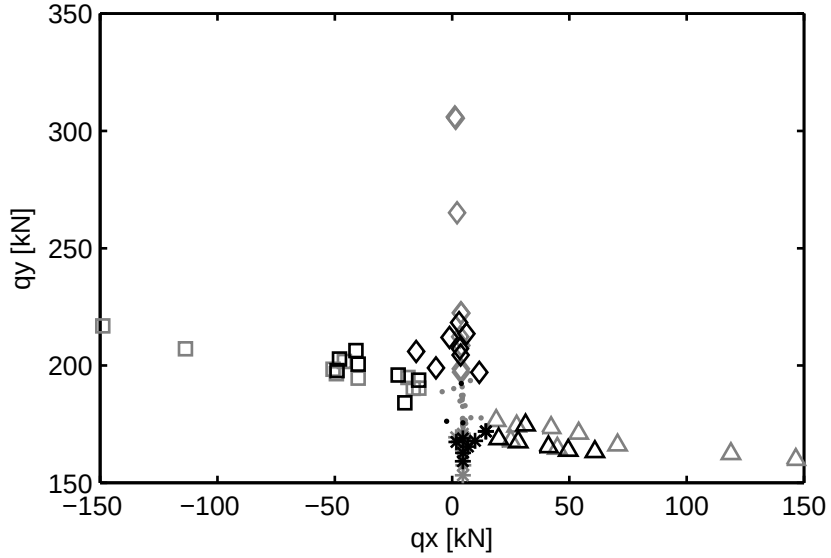
Rys. 6.18. Przykłady testujące — klasy reakcji podporowych, łożysko 1



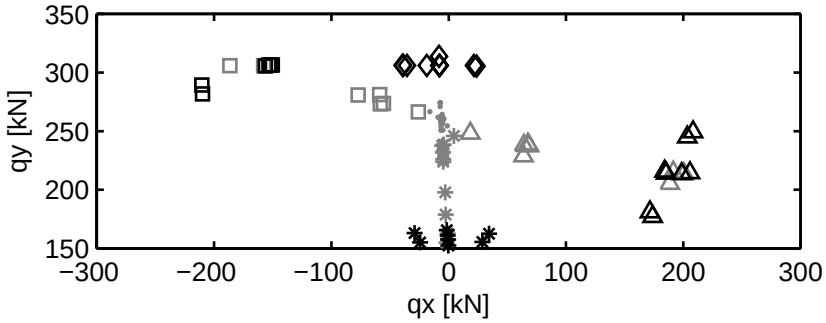
Rys. 6.19. Przykłady testujące — klasy reakcji podporowych, łożysko 2



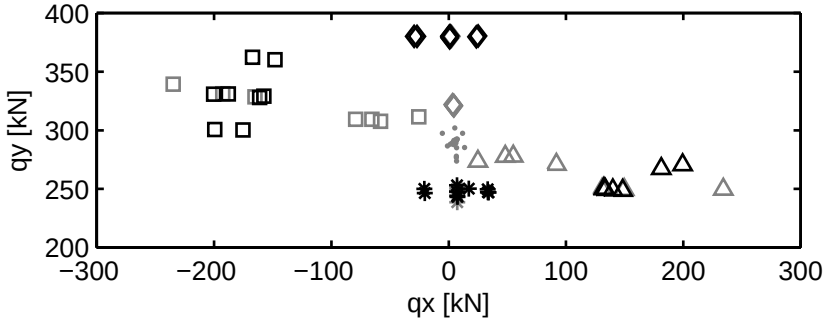
Rys. 6.20. Przykłady testujące — klasy reakcji podporowych, łożysko 3



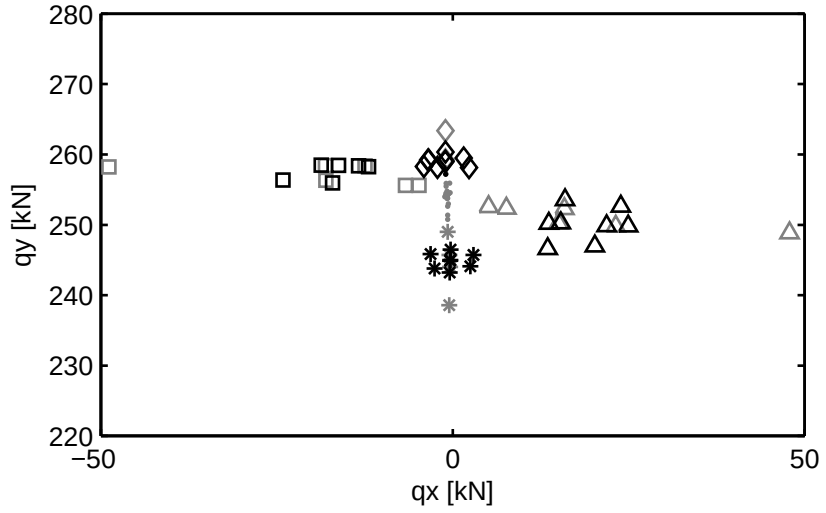
Rys. 6.21. Przykłady testujące — klasy reakcji podporowych, łożysko 4



Rys. 6.22. Przykłady testujące — klasy reakcji podporowych, łożysko 5



Rys. 6.23. Przykłady testujące — klasy reakcji podporowych, łożysko 6



Rys. 6.24. Przykłady testujące — klasy reakcji podporowych, łożysko 7

Rysunki przedstawiające przynależność przykładów testujących do klas reakcji podporowych potwierdzają słuszność stwierdzenia, że na wartości reakcji danej podpory wpływają nie tylko przemieszczenia rozpatrywanej podpory ale również przemieszczenia sąsiednich podpór. Zbiór danych testujących zawiera przypadki przemieszczenia wyłącznie podpory 5 i 6 podczas gdy przekroczenia wartości reakcji poza stan bazowy obserwuje się również na sąsiednich łożyskach.

6.3.3. Wyznaczenie i ocena klasyfikatorów

Kolejne, wymienione w planie badań działania, czyli

- wybór cech relewantnych,
- identyfikacja modeli lokalnych (klasyfikatorów),
- sprawdzenie dopasowania klasyfikatorów do danych trenujących,
- weryfikacja klasyfikatorów na podstawie danych testujących,

przeprowadzono łącznie dla każdego wariantu opracowanej metody (rozdział 6.2.). Wyboru cech relewantnych dokonywano ze zbioru wszystkich 316 cech zawartych w danych symulacyjnych.

Ocenę jakości opracowanych modeli wyznaczano na podstawie uzyskanych dla nich wyników klasyfikacji w dziedzinie cech relewantnych względem wyników klasyfikacji w dziedzinie sygnałów dodatkowych. Wyznaczano następujące liczby:

- O_1 — względna liczba poprawnie zaklasyfikowanych przykładów

$$O_1 := \frac{l_1}{l_w} \cdot 100\%, \tag{6.1}$$

gdzie:
l_1 — liczba poprawnie zaklasyfikowanych przykładów,
l_w — liczba wszystkich klasyfikowanych przykładów,

- O_2 — względna liczba błędów typu „niewykrycie niesprawności"

$$O_2 := \frac{l_2}{l_w} \cdot 100\%, \tag{6.2}$$

gdzie:
l_2 — liczba przykładów, które powinny zostać zaklasyfikowane do jednej z klas: G, D,
L, P, a zostały zaklasyfikowane do klasy S,
l_w — liczba wszystkich klasyfikowanych przykładów,

- O_3 — względna liczba błędów typu „wykrycie nie istniejącej niesprawności"

$$O_3 := \frac{l_3}{l_w} \cdot 100\%, \tag{6.3}$$

gdzie:
l_3 — liczba przykładów, które powinny zostać zaklasyfikowane do klasy S a zostały zaklasyfikowane dojednej z klas: G, D, L, P,
l_w — liczba wszystkich klasyfikowanych przykładów,

- O_4 — względna liczba błędów typu „niepoprawna identyfikacja kierunku przekroczenia wartości nominalnej reakcji"

$$O_4 := \frac{l_4}{l_w} \cdot 100\%, \tag{6.4}$$

gdzie:
l_4 — liczba przykładów zaklasyfikowanych do jednej z klas: G, D, L, P, które powinny
być zaklasyfikowane do innej z tych klas,
l_w — liczba wszystkich klasyfikowanych przykładów.

Ocena O_1 jest ogólną oceną jakości klasyfikatora, większa wartość oceny oznacza lepszą jakość klasyfikatora. Oceny O_2, O_3 oraz O_4 określają względną liczbę błędów klasyfikacji — większa wartość oceny oznacza niższą jakość klasyfikatora; oceny te uszeregowano na podstawie wagi rodzaju błędu w porządku malejącym.

Wyniki oceny wyznaczonych klasyfikatorów (modeli lokalnych) podzielono na dwie grupy:

1. otrzymanych dla wariantów, w których liczba cech była również przedmiotem optymalizacji (warianty 1xx),

2. otrzymanych dla wariantów, w których liczba cech była parametrem (warianty 2xx, 3xx, 4xx).

Dla wariantów, w których optymalizacja prowadzona jest z użyciem algorytmu ewolucyjnego obliczenia powtórzono 20-krotnie, przedstawiono najlepsze z otrzymanych wyników.

Ze względu na dużą liczbę wyników (7 modeli lokalnych, 7 liczności zbioru cech relewantnych, 4 oceny, 36 wariantów metody obliczeń) wstępne porównanie przedstawionych wariantów opracowanej metody pokazane zostanie na podstawie średniej (dla 7 modeli lokalnych) wartości oceny O_1.

W tab. 6.1 przedstawiono wartości średnie (dla modeli lokalnych związanych z łożyskami turbogeneratora) oceny O_1 otrzymanych w wyniku klasyfikacji danych trenujących oraz testujących. Wartości te uzyskano z użyciem klasyfikatorów identyfikowanych wg wariantów 1xx opracowanej metody badań.

W zależności od użytej funkcji oceniającej zmienia się liczność zbioru cech relewantnych. Kilkukrotnie większa liczba cech nie powoduje proporcjonalnie wyższej jakości klasyfikatora, w niektórych przypadkach prowadzi do jej obniżenia.

Tab. 6.1. Wyniki weryfikacji (ocena O_1) wariantów 1xx (ewolucyjna optymalizacja liczby i rodzaju cech; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
oznaczenie metody obliczeń								
111	121	131	112	122	132	113	123	133
liczba cech (średnio)								
92	15	93	92	15	93	92	15	93
wartości oceny O_1 (średnio), dane trenujące								
88,5%	86,0%	84,7%	94,2%	91,2%	93,5%	24,3%	51,4%	26,6%
wartości oceny O_1 (średnio), dane testujące								
60,7%	55,8%	52,7%	52,2%	55,4%	56,7%	24,1%	57,1%	33,0%

W tab. 6.2–6.7 przedstawiono wartości oceny O_1 klasyfikatorów wyznaczonych dla wariantów 2xx, 3xx oraz 4xx proponowanej metody w zależności od liczności zbioru cech relewantnych. Wartości te są średnią ocen modeli lokalnych związanych z łożyskami turbogeneratora. Pogrubioną czcionką zaznaczono najlepsze oceny dla każdego z wariantów opracowanej metody.

Dla parametru „liczność zbioru cech relewantnych" przyjmowano wartości ze zbioru $\{4, 5, ..., 10\}$. Założenie to okazało się słuszne co potwierdzają otrzymane wyniki: otrzymano klasyfikatory o wyższej jakości niż w przypadku dużych liczności zbioru cech relewantnych (tab. 6.1).

Ze względu na małą liczbę dostępnych przykładów uczących klasyfikatory: sztuczna sieć neuronowa oraz algorytm SVM wykazały zbytnie dopasowanie się do danych trenujących co spowodowało ich niską skuteczność w klasyfikacji danych testujących. Warianty opracowanej metody bazujące na tych klasyfikatorach nie będą dalej rozpatrywane.

Analiza otrzymanych wyników badań pozostałych wariantów opracowanej metody prowadzi do następujących spostrzeżeń:

- można wyznaczyć pewną graniczną dla danego wariantu liczność zbioru cech relewantnych, przekroczenie której nie powoduje dalszej poprawy jakości wyznaczanych klasyfikatorów;
- lepsze wyniki otrzymano gdy zbiór cech relewantnych wybierany był jako optymalny zbiór cech (algorytm ewolucyjny, wybieranie z powracaniem) niż jako zbiór optymalnych cech (wybór zadanej liczby najlepszych cech);
- zastosowanie algorytmu ewolucyjnego do poszukiwania zbioru cech relewantnych w większości przypadków daje lepsze rezultaty niż zastosowanie metody wybierania z powracaniem;
- zastosowany prosty klasyfikator ma dobrą zdolność uogólniania: widoczny jest niewielki spadek oceny klasyfikacji danych testujących względem oceny klasyfikacji danych uczących;

Tab. 6.2. Wyniki weryfikacji (ocena O_1) wariantów 2xx, dane trenujące (ewolucyjna optymalizacja zbioru cech o zadanej liczności; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	211	221	231	212	222	232	213	223	233
4	93,2%	84,2%	88,7%	82,5%	68,7%	87,2%	72,4%	77,9%	**77,7%**
5	92,5%	86,0%	90,0%	84,7%	74,4%	87,7%	**81,2%**	**78,2%**	75,4%
6	93,2%	91,2%	88,5%	87,5%	76,4%	88,5%	78,7%	77,7%	74,7%
7	93,7%	90,0%	91,0%	89,5%	80,5%	88,7%	77,4%	77,9%	75,7%
8	94,7%	93,5%	91,7%	91,7%	82,5%	89,5%	79,9%	78,2%	75,2%
9	93,5%	**94,5%**	**93,5%**	92,0%	85,2%	91,0%	51,1%	77,9%	76,9%
10	**95,2%**	93,0%	93,5%	**92,2%**	**87,5%**	**92,0%**	71,7%	77,9%	74,7%

Tab. 6.3. Wyniki weryfikacji (ocena O_1) wariantów 3xx, dane trenujące (wybieranie z powracaniem zadanej liczby cech; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	311	321	331	312	322	332	313	323	333
4	88,5%	83,0%	88,2%	80,7%	68,7%	54,6%	77,7%	76,9%	**64,4%**
5	91,5%	87,7%	90,2%	82,2%	74,2%	70,4%	75,0%	76,9%	58,1%
6	90,0%	91,2%	89,7%	84,7%	76,2%	75,9%	**80,4%**	76,9%	59,1%
7	89,7%	94,0%	91,5%	87,2%	78,2%	79,4%	78,1%	76,7%	45,9%
8	91,5%	93,2%	91,2%	89,2%	81,5%	82,5%	75,4%	**77,4%**	44,6%
9	**92,2%**	93,7%	**92,5%**	91,5%	82,7%	83,0%	72,8%	76,9%	43,9%
10	91,5%	**94,2%**	92,5%	**92,0%**	**85,7%**	**85,0%**	73,7%	76,4%	45,6%

Tab. 6.4. Wyniki weryfikacji (ocena O_1) wariantów 4xx, dane trenujące (wybór zadanej liczby najlepszych cech; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	411	421	431	412	422	432	413	423	433
4	87,0%	83,0%	85,2%	71,4%	68,7%	64,9%	**69,9%**	76,7%	9,5%
5	87,7%	86,7%	85,7%	77,7%	70,4%	55,9%	60,3%	76,9%	**15,8%**
6	87,2%	87,2%	83,2%	68,2%	72,7%	**65,9%**	47,6%	76,4%	12,8%
7	87,0%	88,7%	84,0%	74,7%	75,4%	60,2%	49,4%	76,2%	10,3%
8	88,5%	90,7%	83,5%	76,4%	79,2%	57,4%	49,6%	76,7%	12,3%
9	**90,7%**	**92,0%**	**92,0%**	**85,5%**	82,2%	64,7%	43,8%	**77,2%**	11,5%
10	90,0%	90,5%	86,7%	83,7%	**83,0%**	65,7%	48,2%	77,2%	12,8%

Tab. 6.5. Wyniki weryfikacji (ocena O_1) wariantów 2xx, dane testujące (ewolucyjna optymalizacja zbioru cech o zadanej liczności; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	211	221	231	212	222	232	213	223	233
4	**69,6**%	50,0%	**63,8**%	50,4%	52,2%	54,5%	68,3%	72,3%	71,9%
5	65,6%	58,9%	53,6%	51,3%	52,2%	54,9%	**83,0**%	74,1%	**73,7**%
6	68,3%	59,8%	67,0%	52,7%	50,4%	54,9%	77,2%	74,1%	69,6%
7	64,3%	60,3%	56,7%	**53,1**%	54,9%	54,5%	75,0%	74,6%	71,4%
8	68,3%	60,3%	62,5%	52,2%	55,4%	53,1%	69,6%	75,4%	70,1%
9	66,1%	**62,9**%	60,7%	51,8%	55,4%	57,6%	49,1%	75,4%	71,0%
10	68,8%	59,8%	56,7%	53,1%	**59,4**%	**59,8**%	76,3%	**75,9**%	70,5%

Tab. 6.6. Wyniki weryfikacji (ocena O_1) wariantów 3xx, dane testujące (wybieranie z powracaniem zadanej liczby cech; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	311	321	331	312	322	332	313	323	333
4	62,5%	51,8%	61,6%	50,0%	52,2%	40,6%	**76,2**%	67,9%	**67,0**%
5	**63,4**%	56,7%	60,3%	49,1%	51,3%	52,2%	65,7%	69,6%	54,0%
6	63,4%	62,9%	58,9%	48,2%	49,1%	47,3%	66,7%	69,6%	53,1%
7	61,2%	**64,3**%	51,8%	48,7%	51,8%	49,6%	66,4%	**70,5**%	43,3%
8	57,1%	59,8%	56,7%	48,7%	56,7%	**54,0**%	64,2%	70,5%	43,8%
9	52,7%	55,8%	61,6%	49,6%	56,7%	54,0%	67,4%	70,5%	42,9%
10	61,6%	55,4%	**62,1**%	**54,0**%	**57,6**%	53,6%	67,9%	70,5%	42,9%

Tab. 6.7. Wyniki weryfikacji (ocena O_1) wariantów 4xx, dane testujące (wybór zadanej liczby najlepszych cech; druga cyfra oznacza sposób oceny rozwiązań: 1 — wg zal. (5.1), 2 — wg zal. (5.2), 3 — wg zal. (5.3); trzecia cyfra oznacza rodzaj klasyfikatora)

liczba cech	sztuczna sieć neuronowa			algorytm SVM			prosty klasyfikator		
	oznaczenie metody obliczeń								
	411	421	431	412	422	432	413	423	433
4	62,5%	57,1%	54,9%	46,0%	51,8%	42,0%	**61,0**%	67,9%	16,5%
5	**66,5**%	57,1%	60,3%	50,0%	54,0%	42,0%	58,6%	69,6%	20,5%
6	59,4%	59,4%	58,9%	45,1%	52,7%	43,3%	45,1%	69,2%	23,7%
7	64,3%	51,3%	56,7%	**52,7**%	53,6%	48,2%	45,1%	69,6%	17,9%
8	58,0%	58,0%	57,1%	50,4%	54,5%	42,4%	45,1%	70,1%	24,6%
9	59,4%	**61,6**%	**64,7**%	48,7%	56,3%	**58,9**%	42,9%	**70,5**%	22,8%
10	64,7%	60,7%	54,5%	49,6%	**60,3**%	55,4%	42,1%	70,5%	**25,0**%

- w niektórych przypadkach otrzymano lepszy wynik klasyfikacji danych testujących niż klasyfikacji danych trenujących, jest to spowodowane specyfiką danych uczących.

W tab. 6.8 pokazano średnie wartości (dla modeli lokalnych związanych z łożyskami turbogeneratora) ocen O_1, O_2, O_3 oraz O_4 dla trzech wybranych wariantów opracowanej metody, dla których otrzymano najwyższe wartości oceny O_1. Jako najlepszy wariant opracowanej metody należy wskazać wariant o oznaczeniu 213 (wybór zadanej liczby cech relewantnych z zastosowaniem algorytmu ewolucyjnego, funkcja oceniająca określona na podstawie zgodności najbliższych sąsiadów (5.1), model identyfikowany jako prosty klasyfikator) ze względu na najlepszą ocenę liczby poprawnie zaklasyfikowanych przykładów testujących (ocena O_1) oraz najniższą wartość błędu „niewykrycie niesprawności" (ocena O_2). Duże znaczenie ma również fakt, że wynik ten został osiągnięty przy prawie najmniejszej z rozpatrywanych liczności zbioru cech relewantnych.

Tab. 6.8. Średnie wartości (dla siedmiu modeli lokalnych związanych z łożyskami turbogeneratora) ocen wybranych wariantów opracowanej metody

wariant	213	313	223
liczba cech	5	4	10
ocena O_1	83,0%	76,2%	75,9%
ocena O_2	7,1%	11,8%	11,6%
ocena O_3	2,7%	2,0%	4,0%
ocena O_4	7,2%	10,0%	8,5%

6.4. Analiza wybranego wariantu

W tab. 6.9 pokazano zestawienie ocen modeli lokalnych uzyskanych z użyciem najlepszego wariantu opracowanej metody obliczeń dla zbioru liczącego 5 cech relewantnych. Błąd polegający na niewykryciu niesprawności (ocena O_2) najczęściej jest niewielki lub nie występuje wcale. W przypadku danych testujących błędnie zaklasyfikowane przykłady częściej powodują błędną identyfikację kierunku przekroczenia wartości nominalnej reakcji wykrytej niesprawności (ocena O_4) niż błąd typu „wykrycie nieistniejącej niesprawności" (ocena O_3).

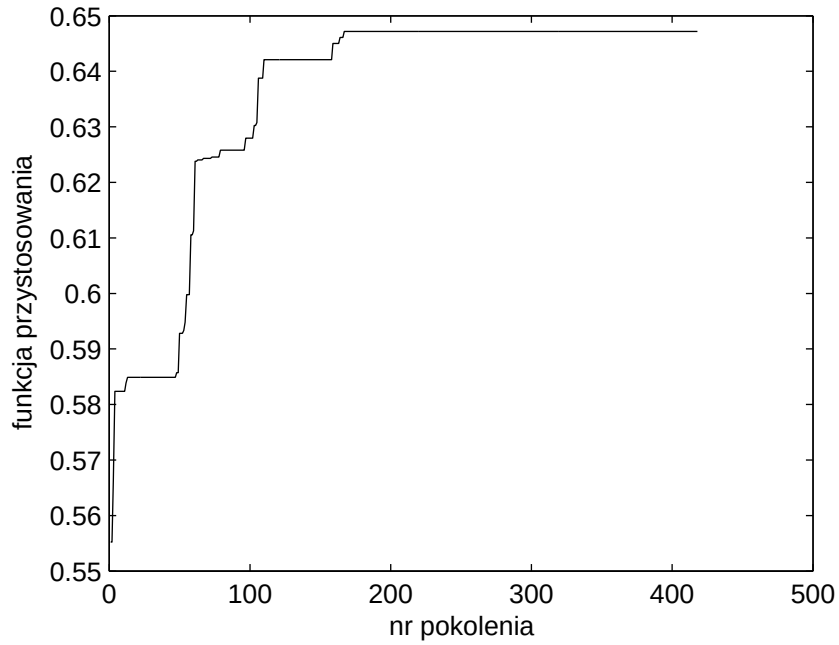
Lepsze wyniki klasyfikacji danych testujących względem wyników klasyfikacji danych trenujących można wytłumaczyć specyfiką zastosowanych danych. W przypadku łożyska pierwszego i drugiego wszystkie przykłady zawarte w danych testujących należą do jednej tylko klasy (rys. 6.18,6.19) (dodatkowo jest to klasa S zawierająca przykłady o nominalnych wartościach reakcji podporowych) co powoduje że dane testujące są 'łatwiejsze' do klasyfikacji niż dane trenujące. Z kolei w przypadku łożyska siódmego dane testujące tworzą bardzo wyraźnie rozdzielone klasy co również ułatwia klasyfikację.

Na rys. 6.25–6.31 pokazano zmiany wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach podczas obliczeń związanych z poszukiwaniem zbiorów

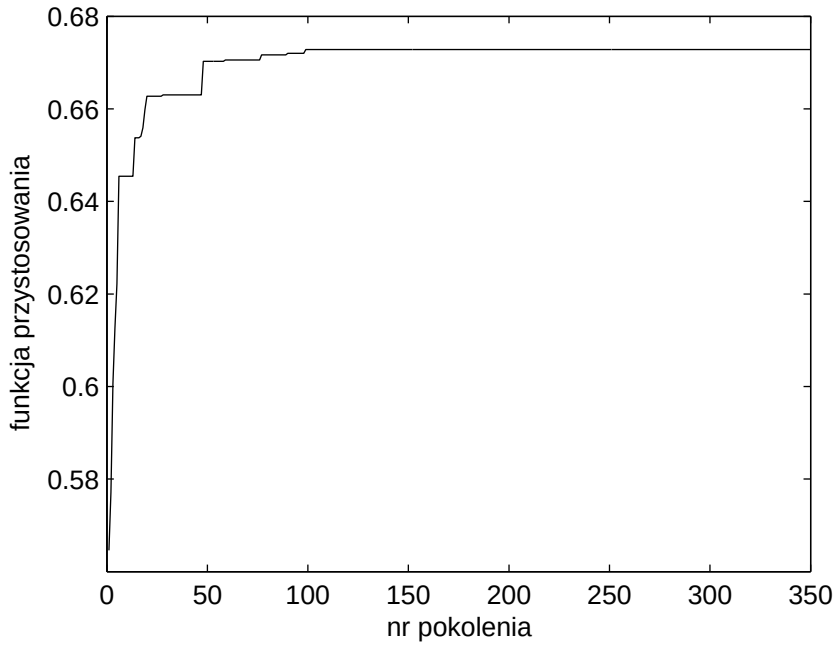
Tab. 6.9. Zestawienie ocen dla wybranego wariantu (213) metody obliczeń

nr łożyska	1	2	3	4	5	6	7
dane trenujące							
ocena O_1	89,5%	93,0%	73,7%	77,2%	87,7%	89,5%	57,9%
ocena O_2	1,8%	0,0%	12,3%	3,5%	0,0%	0,0%	0,0%
ocena O_3	7,0%	1,8%	1,8%	5,3%	1,8%	7,0%	35,1%
ocena O_4	1,8%	5,3%	12,3%	14,0%	10,5%	3,5%	7,0%
dane testujące							
ocena O_1	90,6%	96,9%	71,9%	78,1%	71,9%	87,5%	84,4%
ocena O_2	0,0%	0,0%	25,0%	3,1%	9,4%	0,0%	12,5%
ocena O_3	9,4%	3,1%	3,1%	0,0%	0,0%	0,0%	3,1%
ocena O_4	0,0%	0,0%	0,0%	18,8%	18,8%	12,5%	0,0%

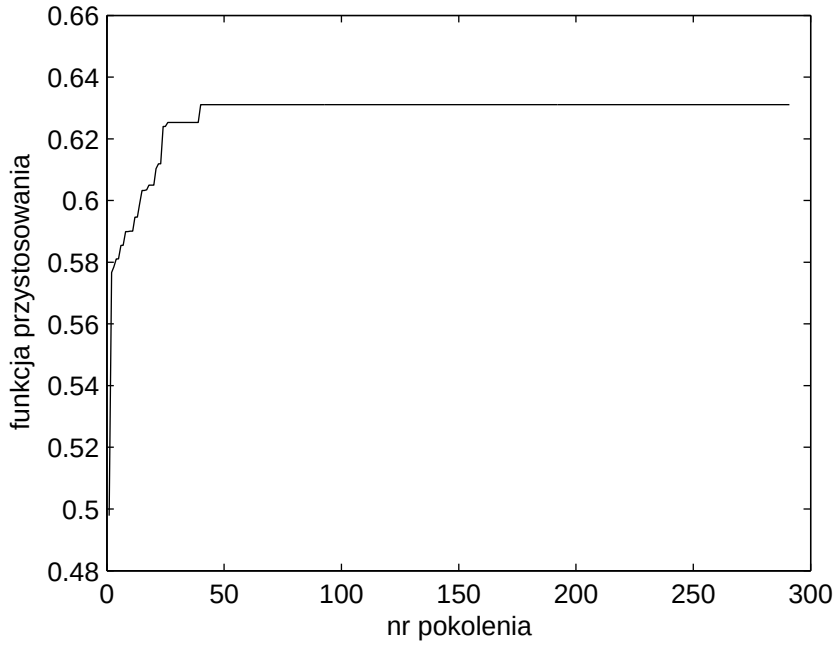
cech relewantnych według najlepszego wariantu opracowanej metody. Niemal za każdym razem obliczenia kończyły się ze względu na kryterium minimalnej szybkości poprawy, jedynie w przypadku łożyska siódmego ze względu na kryterium maksymalnego kosztu. Obliczenia przeprowadzone dla podwojonych wartości progowych przyjętych kryteriów nie przyniosły poprawy wyników — na tej podstawie można wnioskować, że wartości progowe kryteriów zostały przyjęte poprawnie.



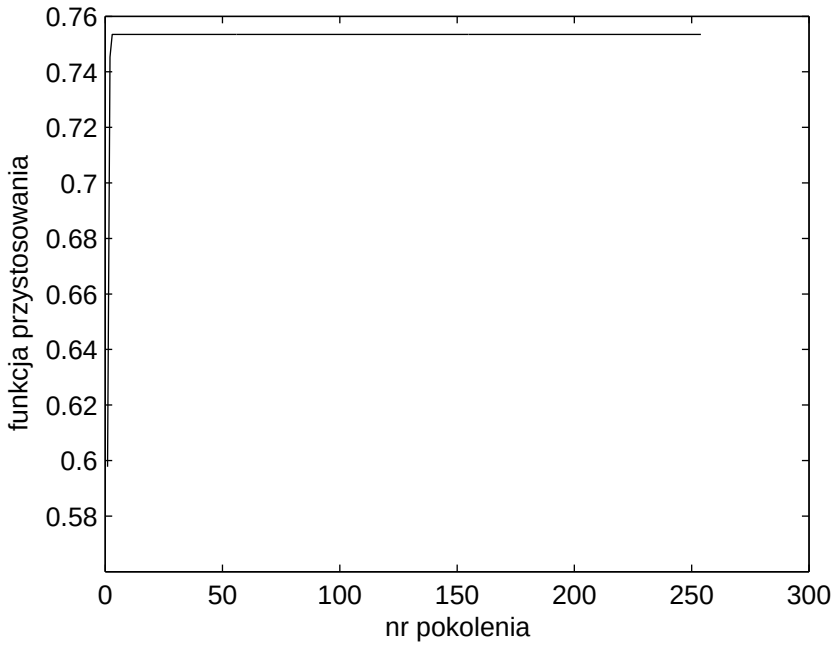
Rys. 6.25. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 1



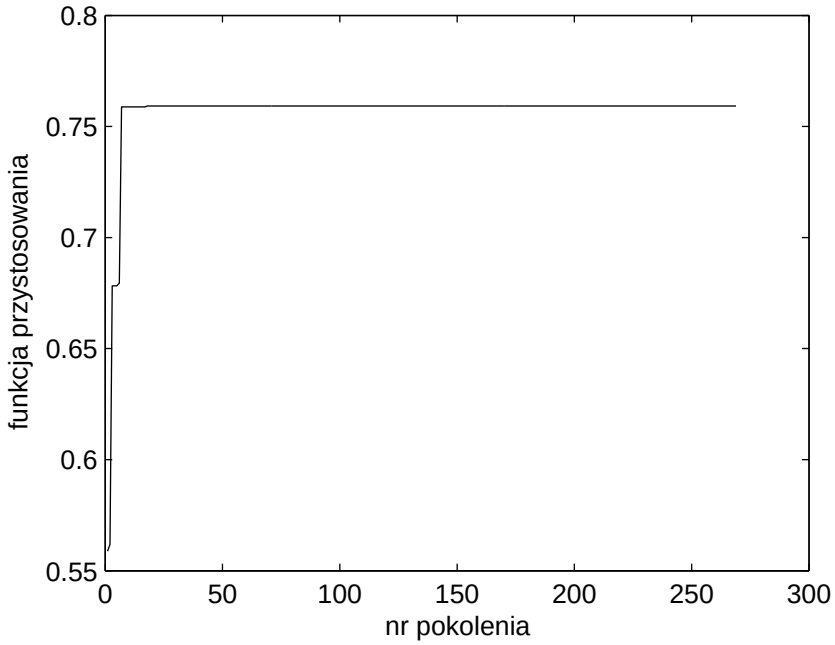
Rys. 6.26. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 2



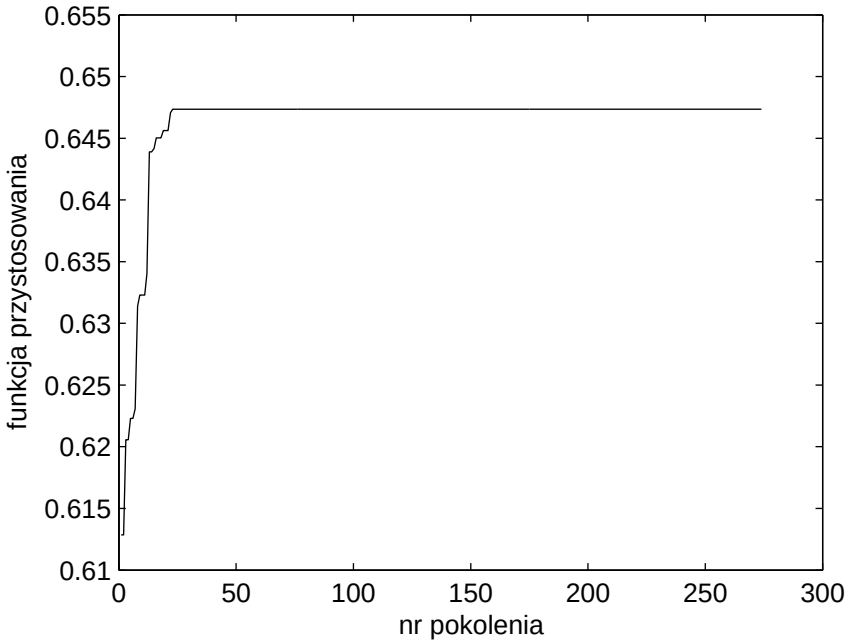
Rys. 6.27. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 3



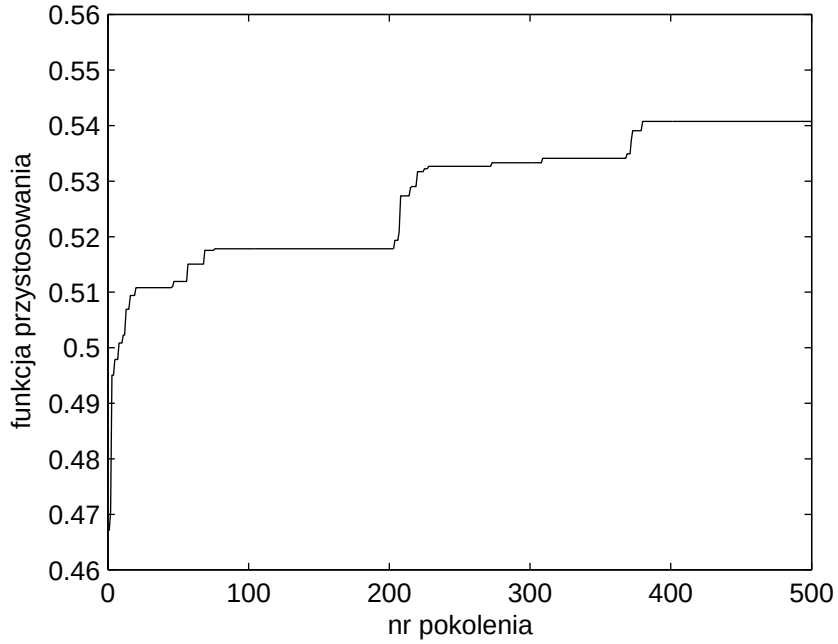
Rys. 6.28. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 4



Rys. 6.29. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 5



Rys. 6.30. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 6



Rys. 6.31. Wartości funkcji przystosowania najlepszego osobnika w kolejnych pokoleniach, łożysko 7

W wyniku obliczeń według najlepszego wariantu opracowanej metody badań otrzymano następujące zbiory cech relewantnych:

- dla łożyska 1
  1. u.lo(1).bca.x.c.a3 (składowa widma amplitudowego o częstotliwości $3X$ sygnału przyspieszeń bezwzględnych czopa w kierunku x)
  2. u.lo(1).bpl.x.c.k (współczynnik kształtu sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku x)
  3. u.lo(1).bca.x.c.a4 (składowa widma amplitudowego o częstotliwości $4X$ sygnału przyspieszeń bezwzględnych czopa w kierunku x)
  4. u.lo(1).bpl.r.c.k (współczynnik kształtu trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
  5. u.lo(1).el.y.c.s (współczynnik asymetrii sygnału przemieszczeń czopa względem panwi w kierunku y)
- dla łożyska 2
  1. u.lo(2).el.y.c.s (współczynnik asymetrii sygnału przemieszczeń czopa względem panwi w kierunku y)
  2. u.lo(2).bcl.y.c.s (współczynnik asymetrii sygnału przemieszczeń bezwzględnych czopa w kierunku y)
  3. u.lo(2).bcl.y.c.f4 (faza składowej widma amplitudowego o częstotliwości $4X$ sygnału przemieszczeń bezwzględnych czopa w kierunku y)
  4. u.lo(2).bca.x.c.a033 (składowa widma amplitudowego o częstotliwości $0.33X$ sygnału przyspieszeń bezwzględnych czopa w kierunku x)
  5. u.lo(2).bpl.r.c.obr1.k2 (obraz nr 2 trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
- dla łożyska 3
  1. u.lo(3).bpl.y.c.rms (wartość skuteczna sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku y)
  2. u.lo(3).bpl.r.c.pp (wartość międzyszczytowa trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
  3. u.lo(3).bcl.rn.c.r0 (liczony w układzie biegunowym promień środka trajektorii niecentralnej przemieszczeń bezwzględnych czopa)
  4. u.lo(3).el.r.c.ppl (wartość szczytowa dodatnia trajektorii centralnej przemieszczeń czopa względem panwi)
  5. u.lo(3).el.r.c.a033 (składowa widma amplitudowego o częstotliwości $0.33X$ trajektorii centralnej przemieszczeń czopa względem panwi)

- dla łożyska 4

    1. u.lo(4).bpl.r.c.f4 (faza składowej widma amplitudowego o częstotliwości $4X$ trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
    2. u.lo(4).el.rn.c.r0 (liczony w układzie biegunowym promień środka trajektorii niecentralnej przemieszczeń czopa względem panwi)
    3. u.lo(4).bcl.y.c.f3 (faza składowej widma amplitudowego o częstotliwości $3X$ sygnału przemieszczeń bezwzględnych czopa w kierunku y)
    4. u.lo(4).bpl.y.c.f4 (faza składowej widma amplitudowego o częstotliwości $4X$ sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku y)
    5. u.lo(4).bca.x.c.a040 (składowa widma amplitudowego o częstotliwości $0.40X$ sygnału przyspieszeń bezwzględnych czopa w kierunku x)

- dla łożyska 5

    1. u.lo(5).bpl.x.c.f3 (faza składowej widma amplitudowego o częstotliwości $3X$ sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku x)
    2. u.lo(5).bpl.r.c.f1 (faza składowej widma amplitudowego o częstotliwości $1X$ trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
    3. u.lo(5).el.rn.c.r0 (liczony w układzie biegunowym promień środka trajektorii niecentralnej przemieszczeń czopa względem panwi)
    4. u.lo(5).bca.r.c.s (współczynnik asymetrii trajektorii centralnej przyspieszeń bezwzględnych czopa sygnału przyspieszeń bezwzględnych czopa)
    5. u.lo(5).bca.x.c.a4 (składowa widma amplitudowego o częstotliwości $4X$ sygnału przyspieszeń bezwzględnych czopa w kierunku x)

- dla łożyska 6

    1. u.lo(6).bcl.y.c.a2 (składowa widma amplitudowego o częstotliwości $2X$ sygnału przemieszczeń bezwzględnych czopa w kierunku y)
    2. u.lo(6).bpl.y.c.pmi (wartość szczytowa ujemna sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku y)
    3. u.lo(6).bpl.r.c.obr1.k2 (obraz nr 2 trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
    4. u.lo(6).bpl.x.c.a2 (składowa widma amplitudowego o częstotliwości $2X$ sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku x)
    5. u.lo(6).el.y.c.a1 (składowa widma amplitudowego o częstotliwości $1X$ sygnału przemieszczeń czopa względem panwi w kierunku y)

- dla łożyska 7

1. u.lo(7).bpl.y.c.a040 (składowa widma amplitudowego o częstotliwości $0.40X$ sygnału przemieszczeń bezwzględnych pokrywy łożyskowej w kierunku y)
2. u.lo(7).bpl.r.c.a2 (składowa widma amplitudowego o częstotliwości $2X$ trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)
3. u.lo(7).el.x.c.a025 (składowa widma amplitudowego o częstotliwości $0.25X$ sygnału przemieszczeń czopa względem panwi w kierunku x)
4. u.lo(7).bcl.r.c.a4 (składowa widma amplitudowego o częstotliwości $4X$ trajektorii centralnej przemieszczeń bezwzględnych czopa)
5. u.lo(7).bpl.r.c.a3 (składowa widma amplitudowego o częstotliwości $3X$ trajektorii centralnej przemieszczeń bezwzględnych pokrywy łożyskowej)

W tab. 6.10— 6.16 pokazano częstość wybierania poszczególnych cech dla 20 powtórzeń obliczeń według najlepszego wariantu opracowanej metody dla wszystkich łożysk turbozespołu.

Tab. 6.10. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 1 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(7) u.lo(1).bpl.r.c.pmi	(5) u.lo(1).bpl.r.c.pp	(5) u.lo(1).bpl.r.c.k
(5) u.lo(1).bcl.y.c.a2	(4) u.lo(1).el.y.c.a2	(4) u.lo(1).bpl.x.c.pp
(4) u.lo(1).bpl.x.c.a040	(4) u.lo(1).bpl.y.c.a1	(3) u.lo(1).bpl.r.c.obr1k3
(3) u.lo(1).bcl.r.c.a4	(3) u.lo(1).bca.x.c.a4	(3) u.lo(1).bca.y.c.a040
(2) u.lo(1).el.x.c.a025	(2) u.lo(1).bpl.y.c.pp	(2) u.lo(1).bpl.y.c.a033
(2) u.lo(1).bpl.r.c.a040	(2) u.lo(1).bcl.x.c.a033	(2) u.lo(1).bcl.r.c.a1
(2) u.lo(1).bca.y.c.k	(1) u.lo(1).el.y.c.s	(1) u.lo(1).el.y.c.a4
(1) u.lo(1).el.r.c.k	(1) u.lo(1).el.r.c.a1	(1) u.lo(1).el.r.c.a4
(1) u.lo(1).bpl.x.c.k	(1) u.lo(1).bpl.x.c.ku	(1) u.lo(1).bpl.x.c.a05
(1) u.lo(1).bpl.x.c.a2	(1) u.lo(1).bpl.y.c.p	(1) u.lo(1).bpl.y.c.ppl
(1) u.lo(1).bpl.y.c.a025	(1) u.lo(1).bpl.y.c.a2	(1) u.lo(1).bpl.r.c.i
(1) u.lo(1).bpl.r.c.ku	(1) u.lo(1).bpl.r.c.a033	(1) u.lo(1).bpl.r.c.a05
(1) u.lo(1).bpl.r.c.a4	(1) u.lo(1).bpl.r.c.f4	(1) u.lo(1).bcl.x.c.s
(1) u.lo(1).bcl.x.c.a025	(1) u.lo(1).bcl.x.c.a4	(1) u.lo(1).bcl.y.c.s
(1) u.lo(1).bcl.r.c.k	(1) u.lo(1).bcl.r.c.obr1k2	(1) u.lo(1).bcl.r.c.obr1k3
(1) u.lo(1).bcl.rn.c.r0	(1) u.lo(1).bcl.rn.c.cosfi0	(1) u.lo(1).bca.x.c.s
(1) u.lo(1).bca.x.c.a040	(1) u.lo(1).bca.y.c.a3	(1) u.lo(1).bca.y.c.f1
(1) u.lo(1).bca.y.c.f4	(1) u.lo(1).bca.r.c.pmi	(1) u.lo(1).bca.r.c.f3
(1) u.lo(1).bca.r.c.obr1k1		

Tab. 6.11. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 2 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(8) u.lo(2).el.rn.c.sinfi0	(7) u.lo(2).bca.r.c.pmi	(6) u.lo(2).el.r.c.obr1k3
(6) u.lo(2).bpl.r.c.obr1k3	(5) u.lo(2).el.r.c.a1	(5) u.lo(2).bpl.y.c.f1
(5) u.lo(2).bpl.r.c.f3	(5) u.lo(2).bcl.r.c.pmi	(3) u.lo(2).el.y.c.s
(3) u.lo(2).bcl.y.c.s	(3) u.lo(2).bca.y.c.s	(2) u.lo(2).el.x.c.a025
(2) u.lo(2).el.x.c.a033	(2) u.lo(2).el.x.c.a3	(2) u.lo(2).el.y.c.a2
(2) u.lo(2).bpl.x.c.ave	(2) u.lo(2).bpl.r.c.ave	(2) u.lo(2).bpl.r.c.obr1k2
(2) u.lo(2).bcl.rn.c.r0	(2) u.lo(2).bcl.rn.c.sinfi0	(2) u.lo(2).bca.x.c.a025
(2) u.lo(2).bca.y.c.a2	(1) u.lo(2).bpl.x.c.rms	(1) u.lo(2).bpl.x.c.p
(1) u.lo(2).bpl.x.c.pmi	(1) u.lo(2).bpl.x.c.s	(1) u.lo(2).bpl.x.c.a05
(1) u.lo(2).bpl.y.c.a05	(1) u.lo(2).bpl.r.c.pmi	(1) u.lo(2).bpl.r.c.ku
(1) u.lo(2).bpl.r.c.f2	(1) u.lo(2).bcl.x.c.a025	(1) u.lo(2).bcl.y.c.pmi
(1) u.lo(2).bcl.y.c.f4	(1) u.lo(2).bcl.r.c.ku	(1) u.lo(2).bcl.r.c.f1
(1) u.lo(2).bcl.rn.c.cosfi0	(1) u.lo(2).bca.x.c.a033	(1) u.lo(2).bca.x.c.a040
(1) u.lo(2).bca.y.c.a025	(1) u.lo(2).bca.r.c.rms	(1) u.lo(2).bca.r.c.i
(1) u.lo(2).bca.r.c.a1	(1) u.lo(2).bca.r.c.a2	

Tab. 6.12. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 3 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(12) u.lo(3).bcl.rn.c.r0	(9) u.lo(3).el.r.c.a1	(9) u.lo(3).bpl.y.c.rms
(9) u.lo(3).bpl.y.c.p	(9) u.lo(3).bpl.r.c.pp	(6) u.lo(3).bpl.y.c.a1
(4) u.lo(3).el.r.c.pmi	(3) u.lo(3).el.r.c.a05	(3) u.lo(3).bpl.x.c.ppl
(3) u.lo(3).bpl.x.c.a2	(3) u.lo(3).bpl.y.c.f3	(2) u.lo(3).el.rn.c.r0
(2) u.lo(3).el.rn.c.cosfi0	(2) u.lo(3).bpl.y.c.ppl	(2) u.lo(3).bpl.y.c.pp
(2) u.lo(3).bpl.r.c.obr1k3	(2) u.lo(3).bcl.r.c.a1	(1) u.lo(3).el.x.c.ave
(1) u.lo(3).el.x.c.pp	(1) u.lo(3).el.x.c.a05	(1) u.lo(3).el.x.c.a040
(1) u.lo(3).el.r.c.ku	(1) u.lo(3).el.r.c.a025	(1) u.lo(3).el.r.c.spp
(1) u.lo(3).el.rn.c.sinfi0	(1) u.lo(3).bpl.x.c.i	(1) u.lo(3).bpl.x.c.f1
(1) u.lo(3).bpl.x.c.f2	(1) u.lo(3).bcl.x.c.a033	(1) u.lo(3).bcl.x.c.f1
(1) u.lo(3).bcl.y.c.a025	(1) u.lo(3).bcl.y.c.a040	(1) u.lo(3).bcl.r.c.pp
(1) u.lo(3).bcl.r.c.ku	(1) u.lo(3).bca.r.c.ku	

Tab. 6.13. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 4 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(20) u.lo(4).el.rn.c.r0	(3) u.lo(4).bpl.y.c.c	(2) u.lo(4).el.xy.c.sto
(2) u.lo(4).el.xy.c.t1	(2) u.lo(4).el.r.c.f3	(2) u.lo(4).bpl.x.c.f4
(2) u.lo(4).bpl.y.c.f1	(2) u.lo(4).bcl.y.c.f3	(2) u.lo(4).bcl.rn.c.sinfi0
(2) u.lo(4).bca.x.c.f2	(2) u.lo(4).bca.r.c.ave	(2) u.lo(4).bca.r.c.ppl
(1) u.lo(4).el.x.c.f4	(1) u.lo(4).el.y.c.c	(1) u.lo(4).el.y.c.i
(1) u.lo(4).el.y.c.f3	(1) u.lo(4).el.y.c.f4	(1) u.lo(4).el.r.c.k
(1) u.lo(4).el.r.c.i	(1) u.lo(4).el.r.c.obr1k2	(1) u.lo(4).el.r.c.obr1k3
(1) u.lo(4).el.rn.c.cosfi0	(1) u.lo(4).bpl.y.c.i	(1) u.lo(4).bpl.y.c.ku
(1) u.lo(4).bpl.y.c.f2	(1) u.lo(4).bpl.y.c.f3	(1) u.lo(4).bpl.y.c.f4
(1) u.lo(4).bpl.xy.c.sto	(1) u.lo(4).bpl.xy.c.t1	(1) u.lo(4).bpl.r.c.k
(1) u.lo(4).bpl.r.c.f1	(1) u.lo(4).bpl.r.c.f4	(1) u.lo(4).bpl.rn.c.sinfi0
(1) u.lo(4).bpl.rn.c.cosfi0	(1) u.lo(4).bcl.x.c.i	(1) u.lo(4).bcl.x.c.ku
(1) u.lo(4).bcl.x.c.f3	(1) u.lo(4).bcl.x.c.f4	(1) u.lo(4).bcl.y.c.c
(1) u.lo(4).bcl.y.c.ku	(1) u.lo(4).bcl.y.c.f4	(1) u.lo(4).bcl.r.c.c
(1) u.lo(4).bcl.r.c.i	(1) u.lo(4).bcl.r.c.f2	(1) u.lo(4).bcl.r.c.f3
(1) u.lo(4).bcl.r.c.obr1k2	(1) u.lo(4).bcl.rn.c.cosfi0	(1) u.lo(4).bca.x.c.pmi
(1) u.lo(4).bca.x.c.pp	(1) u.lo(4).bca.x.c.i	(1) u.lo(4).bca.x.c.s
(1) u.lo(4).bca.x.c.a033	(1) u.lo(4).bca.x.c.a040	(1) u.lo(4).bca.y.c.ave
(1) u.lo(4).bca.y.c.ppl	(1) u.lo(4).bca.y.c.pmi	(1) u.lo(4).bca.y.c.k
(1) u.lo(4).bca.y.c.c	(1) u.lo(4).bca.y.c.i	(1) u.lo(4).bca.y.c.s
(1) u.lo(4).bca.y.c.f2	(1) u.lo(4).bca.y.c.f4	(1) u.lo(4).bca.r.c.i
(1) u.lo(4).bca.r.c.ku	(1) u.lo(4).bca.r.c.a05	(1) u.lo(4).bca.r.c.f1
(1) u.lo(4).bca.r.c.f4	(1) u.lo(4).bca.r.c.smax	(1) u.lo(4).bca.rn.c.sinfi0

Tab. 6.14. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 5 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(20) u.lo(5).el.rn.c.r0	(20) u.lo(5).bca.x.c.a4	(3) u.lo(5).bcl.y.c.i
(2) u.lo(5).el.x.c.ku	(2) u.lo(5).el.x.c.f2	(2) u.lo(5).el.xy.c.sto
(2) u.lo(5).bca.x.c.ave	(2) u.lo(5).bca.y.c.pp	(2) u.lo(5).bca.r.c.i
(2) u.lo(5).bca.r.c.s	(1) u.lo(5).el.y.c.f3	(1) u.lo(5).el.r.c.s
(1) u.lo(5).el.r.c.f2	(1) u.lo(5).el.r.c.f4	(1) u.lo(5).bpl.x.c.ku
(1) u.lo(5).bpl.x.c.f3	(1) u.lo(5).bpl.y.c.k	(1) u.lo(5).bpl.y.c.f2
(1) u.lo(5).bpl.y.c.f3	(1) u.lo(5).bpl.r.c.i	(1) u.lo(5).bpl.r.c.ku
(1) u.lo(5).bpl.r.c.f1	(1) u.lo(5).bpl.r.c.f2	(1) u.lo(5).bpl.r.c.obr1.k1
(1) u.lo(5).bpl.rn.c.cosfi0	(1) u.lo(5).bcl.y.c.c	(1) u.lo(5).bcl.y.c.ku
(1) u.lo(5).bcl.y.c.s	(1) u.lo(5).bcl.xy.c.t1	(1) u.lo(5).bcl.r.c.i
(1) u.lo(5).bcl.r.c.ku	(1) u.lo(5).bcl.r.c.s	(1) u.lo(5).bcl.r.c.f2
(1) u.lo(5).bcl.r.c.obr1.k2	(1) u.lo(5).bcl.rn.c.cosfi0	(1) u.lo(5).bca.x.c.c
(1) u.lo(5).bca.x.c.s	(1) u.lo(5).bca.x.c.a025	(1) u.lo(5).bca.x.c.a040
(1) u.lo(5).bca.x.c.f4	(1) u.lo(5).bca.y.c.ave	(1) u.lo(5).bca.y.c.rms
(1) u.lo(5).bca.y.c.k	(1) u.lo(5).bca.y.c.a2	(1) u.lo(5).bca.y.c.a4
(1) u.lo(5).bca.y.c.a040	(1) u.lo(5).bca.y.c.f1	(1) u.lo(5).bca.y.c.f4
(1) u.lo(5).bca.xy.c.t1	(1) u.lo(5).bca.r.c.ave	(1) u.lo(5).bca.r.c.p
(1) u.lo(5).bca.r.c.pp	(1) u.lo(5).bca.r.c.a4	

Tab. 6.15. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 6 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(20) u.lo(6).bpl.x.c.a2	(20) u.lo(6).bpl.y.c.pmi	(17) u.lo(6).bcl.y.c.a2
(15) u.lo(6).bpl.y.c.pp	(9) u.lo(6).bcl.r.c.a025	(8) u.lo(6).bcl.r.c.a033
(3) u.lo(6).bcl.x.c.a2	(2) u.lo(6).el.y.c.a1	(2) u.lo(6).bpl.x.c.a040
(1) u.lo(6).el.x.c.f1	(1) u.lo(6).el.r.c.rms	(1) u.lo(6).bpl.y.c.f4
(1) u.lo(6).bpl.r.c.pp		

Tab. 6.16. Zestawienie częstości wybierania cech relewantnych w 20 uruchomieniach algorytmu ewolucyjnego, łożysko 7 (oznaczenie cech wg [30])

(liczba wystąpień) identyfikator cechy		
(14) u.lo(7).bcl.x.c.a2	(13) u.lo(7).bpl.r.c.a3	(9) u.lo(7).bpl.y.c.a040
(8) u.lo(7).bpl.y.c.a1	(5) u.lo(7).bpl.r.c.a2	(5) u.lo(7).bcl.r.c.a05
(5) u.lo(7).bcl.r.c.a4	(4) u.lo(7).el.r.c.pmi	(4) u.lo(7).bpl.x.c.pp
(4) u.lo(7).bpl.y.c.a033	(3) u.lo(7).el.x.c.a033	(3) u.lo(7).bpl.y.c.a025
(3) u.lo(7).bcl.y.c.a033	(2) u.lo(7).el.x.c.a025	(2) u.lo(7).bcl.y.c.a025
(1) u.lo(7).el.x.c.pp	(1) u.lo(7).el.x.c.f2	(1) u.lo(7).el.x.c.f3
(1) u.lo(7).el.y.c.rms	(1) u.lo(7).el.y.c.a033	(1) u.lo(7).el.y.c.a1
(1) u.lo(7).el.r.c.a05	(1) u.lo(7).el.r.c.f1	(1) u.lo(7).bpl.y.c.ppl
(1) u.lo(7).bcl.x.c.ave	(1) u.lo(7).bcl.y.c.ave	(1) u.lo(7).bcl.y.c.pmi
(1) u.lo(7).bcl.r.c.i	(1) u.lo(7).bcl.r.c.a025	(1) u.lo(7).bcl.r.c.a033
(1) u.lo(7).bca.x.c.a2		

Można zauważyć, iż:

- w każdym przypadku występują cechy wielokrotnie wybierane jako relewantne w kolejnych przebiegach algorytmu ewolucyjnego (cechy dominujące),
- częstość wybierania tych cech (stopień dominacji) jest większa w przypadku łożysk bardziej obciążonych (łożyska 5, 6), a mniejsza w przypadku łożysk mniej obciążonych (łożyska 1, 2).

Występowanie cech dominujących pokazuje, że mimo iż przebieg obliczeń z użyciem algorytmu ewolucyjnego charakteryzuje duży stopień losowości to otrzymane wyniki nie są przypadkowe. Duża liczba cech wybieranych jako relewantne n«20 razy, świadczy o występowaniu grup cech związanych ze sobą, co powoduje iż wybranie jednej z nich będzie eliminować możliwość wybrania pozostałych w kolejnych krokach. Powtarzanie procedur wyboru z odpowiednim losowym wskazywaniem cech kandydujących do wyboru prowadzi do rożnych wybieranych cech.

6.5. Podsumowanie

Zaproponowana metoda zastosowania algorytmów ewolucyjnych do optymalizacji zbioru cech relewantnych została pozytywnie zweryfikowana. Pokazano również, przez

porównanie wyników licznych wariantów proponowanej metody, że użycie algorytmów ewolucyjnych do poszukiwania optymalnych zbiorów cech daje lepsze rezultaty niż zastosowanie innych algorytmów, szczególnie algorytmów umożliwiających poszukiwanie zbiorów optymalnych cech. Potwierdza to prawdziwość przyjętej tezy głoszącej, że poszukiwanie optymalnego zbioru cech dla identyfikacji odwrotnych modeli diagnostycznych prowadzi do lepszych rezultatów niż poszukiwanie zbioru cech optymalnych.

Wyniki przeprowadzonych badań potwierdziły stwierdzenie o dużym znaczeniu odpowiedniego doboru funkcji oceniającej dla algorytmu ewolucyjnego — użycie różnych funkcji oceniających prowadziło do różnych wyników (w niektórych przypadkach różnice te są znaczące).

Wpływ na otrzymane wyniki miała również liczność zbiorów cech relewantnych. Wyższą jakość wyznaczanych klasyfikatorów uzyskano dla małych liczności zbiorów cech. Tym samym potwierdzono prawdziwość drugiej tezy mówiącej, że nadmierne zwiększanie liczności zbioru cech, na podstawie których zostanie wyznaczony odwrotny model diagnostyczny, prowadzi do spadku jakości wyznaczonego modelu.

Rozdział 7

Podsumowanie i wnioski

7.1. Podsumowanie

W pracy przedstawiono możliwości zastosowania algorytmów ewolucyjnych w identyfikacji odwrotnych modeli diagnostycznych na podstawie przykładów. Omówiono zagadnienia związane z poszukiwaniem rozwiązania optymalnego z użyciem algorytmów ewolucyjnych: sposoby kodowania rozwiązań w chromosomie, rodzaje operatorów genetycznych, warunki zakończenia działania algorytmu ewolucyjnego, podkreślono duże znaczenie poprawnego zdefiniowania funkcji oceniającej. Przedstawiono metody identyfikacji modeli diagnostycznych wskazując na duże znaczenie modeli odwrotnych. Zwrócono uwagę na konieczność optymalizacji procesu identyfikacji modeli odwrotnych podkreślając możliwość zastosowania w tym celu algorytmów ewolucyjnych.

Przedstawiono przykład optymalizacji modeli lokalnych stanowiących pierwszy stopień dwustopniowego modelu odwrotnego służącego do rozpoznawania stanu technicznego turbogeneratora w zakresie przemieszczeń podpór łożyskowych. Opisano sposób optymalizacji poprzez wybór optymalnego zbioru cech sygnałów diagnostycznych będących podstawą wyznaczania klasyfikatorów stanowiących modele lokalne. Zbadano wpływ metod wyboru zbioru cech oraz jego liczności na jakość otrzymanych wyników. Badania przeprowadzono z użyciem danych uczących pochodzących z badań symulacyjnych prowadzonych w Instytucie Maszyn Przepływowych Polskiej Akademii Nauk w Gdańsku. Na podstawie otrzymanych wyników badań sformułowano wnioski.

7.2. Wniosek ogólny

Otrzymane wyniki przeprowadzonych badań potwierdzają słuszność postawionych w pracy tez. Stwierdzono, że poszukiwanie optymalnego zbioru cech dla identyfikacji odwrotnych modeli diagnostycznych prowadzi do lepszych rezultatów niż poszukiwanie zbioru cech optymalnych. Ponadto stwierdzono, że nadmierne zwiększanie liczności zbioru cech, na podstawie których zostanie wyznaczony odwrotny model diagnostyczny, prowadzi do spadku jakości wyznaczonego modelu.

7.3. Wnioski szczegółowe

1. Optymalizowany model odwrotny był częścią modelu dwustopniowego. Wyodrębnione stopnie były połączone poprzez sygnały pomocnicze. Rozwiązanie to pozwoliło na znaczne uproszczenie struktury modeli składowych i sposobu ich identyfikacji.

2. Użycie klas obciążeń łożysk jako sygnałów pomocniczych pozwoliło na uwzględnienie wpływu stanu jednego łożyska na łożyska sąsiednie. Pozwoliło również na uwzględnienie dodatkowej wiedzy o modelowanych relacjach w identyfikowanych klasyfikatorach.

3. Optymalizacja zbiorów cech połączona z optymalizacją ich liczności prowadziła do zbiorów o dużej liczności (liczba cech relewantnych stanowiła nawet ok. 30% liczby wszystkich rozpatrywanych cech). Klasyfikatory trenowane na podstawie otrzymanych zbiorów charakteryzowały się niską skutecznością klasyfikacji przykładów testujących (poniżej 50%).

4. Spośród dwóch zastosowanych metod poszukiwania optymalnego zbioru cech lepsze wyniki osiągnięto z użyciem algorytmu ewolucyjnego.

5. Przedział zmienności liczności zbiorów cech relewantnych został przyjęty prawidłowo — najlepsze wyniki osiągnięto dla zbiorów o licznościach z wnętrza tego przedziału.

7.4. Kierunki dalszych badań

Analiza otrzymanych wyników pozwala na wskazanie kierunków dalszych badań:

1. O jakości modeli identyfikowanych wg opracowanej metody decyduje zbiór cech relewantnych oraz sposób budowania klasyfikatora. Na jakość zbioru cech relewantnych decydujący wpływ ma użyta w algorytmie ewolucyjnym funkcja przystosowania. Dalsze badania mające na celu poprawę jakości identyfikowanych modeli lokalnych należałoby prowadzić w dwóch kierunkach:

   - poszukiwania innych funkcji przystosowania,
   - poszukiwania innych algorytmów klasyfikacji.

2. Możliwa jest dalsza redukcja wymiarów zadania identyfikacji modeli lokalnych poprzez analizę wyników wielu uruchomień algorytmu ewolucyjnego i poszukiwanie cech dominujących w otrzymywanych zbiorach cech relewantnych.

Bibliografia

[1] Arabas J.: *Algorytmy ewolucyjne: przegląd tematyki*. Proceedings of the Symposium on Artificial Intelligence in Mechanics and Mechanical Engineering AI-MECH 2000, Gliwice 2000, s. 3–23.

[2] Arabas J.: *Wykłady z algorytmów genetycznych*. WNT, Warszawa, 2001.

[3] Arabas J., Michalewicz Z., Mulawka J.: *GAVaPS — a Genetic Algorithm with Varying Population Size*. Proceedings of the First IEEE International Conference on Evolutionary Computation, Orlando 1994, s. 73–78.

[4] Bagley J.D.: *The behavior of adaptive systems which employ genetic and correlation algorithms*. Rozprawa doktorska, University of Michigan, Ann Arbor, 1967.

[5] Baker J.E.: *Adaptive selection methods for genetic algorithms*. Proceedings of an International Conference on Genetic Algorithms and Their Applications, 1985, s. 101–111.

[6] Barszcz T., Czop P., Uhl T.: *Nieparametryczna metoda estymacji w diagnostyce członów wykonawczych systemu sterowania turbiną*. Materiały VI Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'03, Władysławowo 2003, s. 435–440.

[7] Barszcz T., Uhl T., Rumak T.: *Selected applications of modern diagnostics techniques in ALSTOM POWER*. Materiały II Międzynarodowego Kongresu Diagnostyki Technicznej Diagnostyka 2000, Warszawa 2000.

[8] Bäck T.: *Evolutionary algorithms in theory and practice*. Oxford University Press, New York, 1996.

[9] Bäck T., Fogel D.B., Michalewicz Z. (ed.): *Handbook of evolutionary computation*. Oxford University Press, New York, 1996.

[10] Brindle A.: *Genetic Algorithms for function optimization*. Rozprawa doktorska, University of Alberta, Edmonton, 1981.

[11] Burczyński T., Długosz A., Kuś W.: *Improvements of distributed evolutionary algorithm in optimization of nonlinear structures*. Proceedings of the Symposium on Methods of Artificial Intelligence AI-METH 2004, Gliwice 2004.

[12] Burczyński T., Kuś W.: *Shape optimization of heat radiators using parallel evolutionary algorithms*. Proceedings of the Symposium on Methods of Artificial Intelligence AI-METH 2003, Gliwice 2003.

[13] Burges C.: *A Tutorial on Support Vector Machines for Pattern Recognition*. Data Mining and Knowledge Discovery, 2/1998, s. 121–167.

[14] Cantú-Paz E.: *Order statistics and selection methods of evolutionary algorithms*. Information Processing Letters, 82/2002, s. 15–22.

[15] Cavicchio D.J.: *Adaptive search using simulated evolution*. Rozprawa doktorska, University of Michigan, Ann Arbor, 1970.

[16] Cempel Cz.: *Podstawy wibroakustycznej diagnostyki maszyn*. WNT, Warszawa, 1982.

[17] Cempel Cz.: *Diagnostyka wibroakustyczna maszyn*. PWN, Warszawa, 1989.

[18] Chen J., Lui G.P., Patton R.J.: *Optimal residual design for fault diagnosis using multiobjective optimization and genetic algorithms*. Int. J. Syst. Sci., 27/1997, s. 567–576.

[19] Cholewa W.: *Metoda diagnozowania maszyn z zastosowaniem zbiorów rozmytych*. Politechnika Śląska, Gliwice, 1983. Zeszyty Naukowe – Mechanika z. 79.

[20] Cholewa W.: *Projekt szkieletowego systemu doradczego przeznaczonego do dozorowania eksploatacyjnego i diagnostyki turbozespołu*. Materiały Ogólnopolskiego Seminarium Wibrodiagnostyka turbozespołów energetycznych, Gdańsk 1994, s. 25–49.

[21] Cholewa W.: *Dynamiczne systemy doradcze w diagnostyce technicznej*. Materiały Kongresu Diagnostyki Technicznej, Gdańsk 1996, s. 57–84.

[22] Cholewa W.: *Diagnostyka symulacyjna*. Problemy Eksploatacji, 4/1997, s. 13–23.

[23] Cholewa W.: *Real-Time Expert Systems for Technical Diagnostics*. Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, tom 6, Berlin 1997, s. 715–720.

[24] Cholewa W.: *Dynamiczne systemy doradcze*. Proceedings of the Symposium on Artificial Intelligence in Mechanics and Mechanical Engineering AI-MECH 2000, Gliwice 2000, s. 37–54.

[25] Cholewa W.: *Wykład z przedmiotu metody heurystyczne*. Materiały niepublikowane, Gliwice, 2001.

[26] Cholewa W.: *Systemy doradcze w diagnostyce technicznej*. Rozdział 15 [w:] Cholewa W., Korbicz J., Kościelny J.M., Kowalczuk Z. (red.): *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania.* WNT, Warszawa, 2002.

[27] Cholewa W.: *Dekompozycja modeli diagnostycznych*. Materiały VII Krajowej Konferencji Naukowo-Technicznej Diagnostyka Procesów Przemysłowych, Rajgród 2005.

[28] Cholewa W.: *Diagnostyczne modele obiektów*. Rozdział 2 [w:] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Wydawnictwo IMP PAN, Gdańsk, 2005.

[29] Cholewa W.: *System doradczy K015*. Rozdział 27 [w:] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Wydawnictwo IMP PAN, Gdańsk, 2005.

[30] Cholewa W.: *Zbiory danych*. Rozdział 25 [w:] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Wydawnictwo IMP PAN, Gdańsk, 2005.

[31] Cholewa W., Kaźmierczak J.: *Diagnostyka techniczna maszyn. Przetwarzanie cech sygnałów*. Politechnika Śląska, Gliwice, 1992. Skrypty Uczelniane nr 1693.

[32] Cholewa W., Kiciński J.: *Metody identyfikacji reguł dla diagnostycznych systemów doradczych*. XXXV Sympozjon Modelowanie w Mechanice t.1, str. 57–66, Zeszyty Naukowe Katedry Mechaniki Technicznej, Gliwice, 1996.

[33] Cholewa W., Kiciński J. (red.): *Diagnostyka techniczna. Odwrotne modele diagnostyczne*. Wydawnictwo Politechniki Śląskiej, Gliwice, 1997.

[34] Cholewa W., Kiciński J. (red.): *DT200-1. System diagnostyczny dla turbozespołów energetycznych o mocy 200MW*. KPKM Pol.Śl. i IMP PAN, Gdańsk, 1998. PBZ 038-06.

[35] Cholewa W., Kiciński J. (red.): *Diagnostyka techniczna. Metody odwracania nieliniowych modeli obiektów*. Katedra Podstaw Konstrukcji Maszyn, Gliwice, 2001. Zeszyt 120.

[36] Cholewa W., Korbicz J., Kościelny J.M., Kowalczuk Z. (red.): *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania*. WNT, Warszawa, 2002.

[37] Cholewa W., Pedrycz W.: *Systemy doradcze*. Politechnika Śląska, Gliwice, 1987. Skrypty Uczelniane nr 1447.

[38] Cholewa W., Urbanek G.: *Optimization of examples for identification of inverse diagnostic models*. Proceedings of the Symposium on Artificial Intelligence in Mechanics and Mechanical Engineering AI-MECH 2001, Gliwice 2001.

[39] Cholewa W., Urbanek G.: *Optymalizacja ewolucyjna zbioru uwzględnianych cech*. Rozdział 26.4 [w:] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Wydawnictwo IMP PAN, Gdańsk, 2005.

[40] Cholewa W., Urbanek G., Wojtusik J.: *Grupowanie trajektorii ruchu środka czopa*. Materiały VII Krajowej Konferencji Naukowo-Technicznej Diagnostyka Procesów Przemysłowych, Rajgród 2005, s. 231–233.

[41] Cholewa W., Urbanek G., Wojtusik J.: *Modele wielostopniowe. Przykład*. Rozdział 27.5 [w:] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Wydawnictwo IMP PAN, Gdańsk, 2005.

[42] Cholewa W., White M.F.: *Inverse modeling in rotordynamics for identification of unbalance distribution*. Machine Vibration, 2/1993, s. 157–167.

[43] Cholewa W., White M.F.: *Application of neural networks for inverting of vibration models*. Politechnika Śląska, Gliwice, 1994. Zeszyty naukowe – Mechanika z. 116.

[44] Cholewa W., White M.F.: *Inverse vibration modelling for diagnosis of unbalance in machinery*. Proceedings of the 15th Internation Congress on Acoustics, Trondheim 1995, s. 503–506.

[45] Czop P.: *Linear and nonlinear diagnostic model of rotating machine under transient operating conditions — A comparison*. Proceedings of the Symposium on Artificial Intelligence in Mechanics and Mechanical Engineering AI-MECH 2001.

[46] Davidson J., Savic D., Walters G.: *Symbolic and numerical regression: experiments and applications*. Information Sciences, 150/2003, s. 95–117.

[47] De Jong K.A.: *An analysis of the behavior of a class of genetic adaptive systems*. Rozprawa doktorska, University of Michigan, Ann Arbor, 1975.

[48] Demuth H., Beale M.: *Neural Network Toolbox. User's Guide*. The Math Works Inc., 1998.

[49] Derksen S., Keselman H.: *Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables*. British Journal of Mathematical and Stastical, 45/1992, s. 265–282.

[50] Dietrych J.: *System i konstrukcja*. WNT, Warszawa, 1978.

[51] Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R.: *Biocybernetyka i inżynieria biomedyczna. Sieci neuronowe*. Akademicka Oficyna Wydawnicza Exit, Warszawa, 2000.

[52] Duda R.O., Hart P.E.: *Pattern Classification and Scene Analysis*. John Wiley & Sons Inc., 1973.

[53] Dziecioł E., Szczepanowski A.: *Zastosowanie mikrofalowego czujnika ruchu w diagnozowaniu łopatek turbin*. Materiały V Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'01, Łagów Lubuski 2001, s. 401–404.

[54] Eiben A.E., Rudolph G.: *Theory of evolutionary algorithms: a bird's eye view*. Theoretical Computer Science, 229/1999, s. 3–9.

[55] Electronic Statistic Textbook: *http://www.ststsoft.com/textbook/stathome.html*. 2004.

[56] Fogel L.J., Owens A.J., Walsh M.J.: *Artificial Inteligence Through Simulated Evolution*. John Wiley, Chichester, 1966.

[57] Genetic Algorithms: *http://www.econ.iastate.edu/tesfatsi/holland.GAIntro.htm*.

[58] Genetic Algorithms, Evolutionary Computation: *http://www.talkorigins.org/faqs/genalg/genalg.html*.

[59] Goldberg D.E.: *Computer-aided gas pipeline operation using genetic algorithms and rule learning*. Rozprawa doktorska, University of Michigan, Ann Arbor, 1983.

[60] Goldberg D.E.: *Algorytmy genetyczne w zastosowaniach*. WNT, Warszawa, 1995.

[61] Goldberg D.E.: *Algorytmy genetyczne i ich zastosowania*. WNT, Warszawa, 1998.

[62] Goldberg D.E., Samtani M.P.: *Engineering optimization via genetic algorithm*. Proceedings of the Ninth Conference on Electronic Computation, 1986, s. 471–482.

[63] Gryboś R.: *Dynamika maszyn wirnikowych*. IPPT PAN, Warszawa, 1994.

[64] Holland J.H.: *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, 1975.

[65] Hollstien R.B.: *Artificial genetic adaptation in computer control systems*. Rozprawa doktorska, University of Michigan.

[66] Hughes N.P., Lowe D.: *Artefactual Structure from Least Squares Multidimensional Scaling*.

[67] Kacprzyński B.: *Podstawy eksperymentów, podstawy matematyczne*. WNT, Warszawa, 1974.

[68] Kiciński J.: *Symptomy diagnostyczne turbozespołów energetycznych określone metodą komputerowej i analogowej symulacji ich defektów*. Materiały Ogólnopolskiego Seminarium Wibrodiagnostyka turbozespołów energetycznych, Gdańsk 1994, s. 51–92.

[69] Kiciński J.: *Funkcje modeli symulacyjnych w diagnostyce technicznej*. Materiały III Krajowej Konferencji Diagnostyka techniczna urządzeń i systemów DIAG'95, Szczyrk 1995, s. 10–13.

[70] Kiciński J.: *Metody symulacyjnego pozyskiwania relacji diagnostycznych i możliwości ich wykorzystania w budowie systemu nadzoru drganiowego nowej generacji*. Przegląd Mechaniczny, 1–2/1995, s. 5–10.

[71] Kiciński J.: *Model–Based Diagnostics in Rotating Machinery*. Proceedings of the 9th IMEKO TC-10 International Conference on Technical Diagnostics, Wrocław 1999, s. 56–67.

[72] Kiciński J.: *Numerical modelling of selected defects in rotating machinery*. Proceedings of the 5th International Conference Acoustical and Vibratory Surveillence Methods and Diagnostic Technics, Senlis 2004.

[73] Kiciński J. (red.): *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. IMP PAN, Gdańsk, 2005.

[74] Kiciński J. (red.): *Symulacja defektów i budowa relacji diagnostycznych modelu numerycznego*. Katedra Podstaw Konstrukcji Maszyn, Gliwice, 2005. Zeszyt 125.

[75] Klimek A.: *Metody doskonalenia diagnostycznych modeli odwrotnych*. Materiały III Krajowej Konferencji Naukowo Technicznej Diagnostyka Procesów Przemysłowych, Jurata 1998, s. 79–84.

[76] Klimek A.: *Metody doskonalenia odwrotnych modeli diagnostycznych*. Politechnika Śląska, Gliwice, 1999. Zeszyty naukowe — Mechanika z. 134.

[77] Korbicz J., Obuchowicz A.: *Metody ewolucyjne w projektowaniu sysytemów diagnstycznych*. Rozdział 8 [w:] Cholewa W., Korbicz J., Kościelny J.M., Kowalczuk Z. (red.): *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania* WNT, Warszawa, 2002.

[78] Korbicz J., Obuchowicz A., Witczak M.: *Design of non–linear state observers using genetic programming*. Materiały III Krajowej Konferencji Algorytmy Ewolucyjne i Optymalizacja Globalna, Potok Złoty 1999, s. 345–352.

[79] Korbicz J., Witczak M.: *Programowanie genetyczne w diagnostyce uszkodzeń i identyfikacji nieliniowych systemów dynamicznych*. Rozdział 12 [w:] Cholewa W.,

Korbicz J., Kościelny J.M., Kowalczuk Z. (red.): *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania* WNT, Warszawa, 2002.

[80] Kostka P.: *Metody klasyfikacji postaci kinetostatycznych linii wałów maszyn wirnikowych*. Katedra Podstaw Konstrukcji Maszyn, Gliwice, 2001. Zeszyt 121.

[81] Kotowski J., Radkowski S.: *Wykorzystanie analizy korelacyjnej w diagnostyce turbin*. Materiały II Międzynarodowego Kongresu Diagnostyki Technicznej Diagnostyka 2000, Warszawa 2000.

[82] Kowalczuk Z., Białaszewski T.: *Algorytmy genetyczne w wielokryterialnej optymalizacji obserwatorów detekcyjnych*. Rozdział 13 [w:] Cholewa W., Korbicz J., Kościelny J.M., Kowalczuk Z. (red.): *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania* WNT, Warszawa, 2002.

[83] Koza J.R.: *Evolving a Computer Program to Generate Random Numbers Using the Genetic Programming Paradigm*. Proceedings of the Fourth Internation Conference on Genetic Algorithms, San Mateo 1991, s. 37–44.

[84] Koza J.R.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, 1992.

[85] Kozdrowski S.: *Zastosowanie algorytmów ewolucyjnych do projektowania sieci telekomunikacyjnych*. Rozprawa doktorska, Politechnika Warszawska, Warszawa, 1999.

[86] Kozieł S.: *Algorytmy ewolucyjne i ich zastosowania do optymalizacji i modelowania analogowych układów elektronicznych*. Rozprawa doktorska, Politechnika Gdańska, Gdańsk, 2000.

[87] Kubota N., Fukuda T.: *Genetic algorithms with age structure*. Soft Computing, 1/1997, s. 155–161.

[88] Mańczak K.: *Technika planowania eksperymentu*. WNT, Warszawa, 1976.

[89] Michalewicz Z.: *Algorytmy genetyczne + struktury danych = programy ewolucyjne*. WNT, Warszawa, 1999.

[90] Minga A.K.: *Genetic Algorithms in Aerospace Design*. 1986.

[91] Mitchell M., Taylor C.E.: *Evolutionary Computation: An Overview*. Annual Review of Ecology and Systematics, 20/1999, s. 593–616.

[92] Moczulski W.: *Metoda wibroakustyczna badań maszyn wirnikowych w warunkach rozruchu lub zatrzymania*. Rozprawa doktorska, Politechnika Śląska, Gliwice, 1984.

[93] Moczulski W.: *Metody pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*. Politechnika Śląska, Gliwice, 1997. Zeszyty Naukowe – Mechanika z. 130.

[94] Moczulski W.: *Pozyskiwanie wiedzy deklaratywnej i proceduralnej*. Proceedings of the Symposium on Artificial Intelligence in Mechanics and Mechanical Engineering AI-MECH 2000, Gliwice 2000, s. 37–54.

[95] Moczulski W.: *Diagnostyka techniczna. Metody pozyskiwania wiedzy*. Politechnika Śląska, Gliwice, 2002. Monografie – Budowa i Eksploatacja Mazyn.

[96] Moczulski W.: *Metodyka heurystycznego modelowania obiektów i procesów dynamicznych w diagnostyce i sterowaniu*. Materiały VII Krajowej Konferencji Naukowo-Technicznej Diagnostyka Procesów Przemysłowych, Rajgród 2005, s. 39–41.

[97] Polański Z.: *Metodyka badań doświadczalnych*. Wydawnictwo Politechniki Krakowskiej, Kraków, 1981.

[98] Radkowski S.: *Wykorzystanie sygnału wibroakustycznego w diagnozowaniu wczesnych faz uszkodzen*. Materiały VI Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'03, Władysławowo 2003, s. 33–42.

[99] Rybczyński J.: *Przemieszczenia termiczne i montażowe podpór łożyskowych*. IMP PAN, 2004.

[100] Schmitt L.M.: *Theory of genetic algorithms*. Theoretical Computer Science, 259/2001, s. 2–60.

[101] Schmitt L.M.: *Theory ofGenetic Algorithms II: models for genetic operators over the string-tensor representation of populations and convergence to global optima for arbitrary fitness function under scaling*. Theoretical Computer Science, 310/2004, s. 181–231.

[102] Schwefel H.P.: *Evolution and optimum seeking*. Wiley, Chichester, 1995.

[103] Sneath P.H.A., Sokal R.R.: *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. Freeman, San Francisco, 1973.

[104] Stork D.G., Yom-Tov E.: *Computer Manual in MATLAB to accompany Pattern Classification*. John Wiley & Sons, New Jersey, 2004.

[105] Syfert M., Wnuk P., Kościelny J.M.: *System zaawansowanego monitorowania i diagnostyki „AMandD"*. Materiały VII Krajowej Konferencji Naukowo-Technicznej Diagnostyka Procesów Przemysłowych, Rajgród 2005, s. 157–159.

[106] Szulim R.: *Koncepcja systemu doradczego do wspomagania obsługi w prowadzeniu złożonego procesu przemysłowego*. Materiały VI Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'03, Władysławowo 2003.

[107] Tadeusiewicz R.: *Sieci neuronowe*. Akademicka Oficyna Wydawnicza RM, Warszawa, 1993.

[108] The Hitch-Hiker's Guide to Evolutionary Computation: *http://www.cs.bham.ac.uk/Mirrors/ftp.de.uu.net/EC/clife/www/location.htm*. 2005.

[109] Timofiejczuk A.: *Metoda badania maszyn wirnikowych w warunkach rozruchu, rozbiegu i wybiegu*. Politechnika Śląska, Gliwice, 1999. Zeszyty Naukowe – Mechanika z. 133.

[110] Urbanek G.: *Identyfikacja odwrotnych modeli diagnostycznych na podstawie numerycznych modeli złożonych obiektów rzeczywistych*. Materiały V Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'01, Łagów Lubuski 2001.

[111] Urbanek G.: *Application of genetic algorithm for direct controlling of simulation research*. Proceedings of the Symposium on Methods of Artificial Intelligence AI-METH 2002, Gliwice 2002.

[112] Urbanek G.: *Czynniki wpływające na jakość i koszt rozwiązania w identyfikacji modeli odwrotnych z zastosowaniem algorytmu ewolucyjnego*. Materiały VI Krajowej Konferencji Naukowo–Technicznej Diagnostyka Procesów Przemysłowych DPP'03, Władysławowo 2003.

[113] Urbanek G.: *Factors having an effect on quality and cost of solution in identyfication of inverse models with the application of evolutionary algorithms*. Proceedings of the Symposium on Methods of Artificial Intelligence AI-METH 2003, Gliwice 2003.

[114] Urbanek G.: *Algorytmy ewolucyjne w identyfikacji odwrotnych modeli diagnostycznych*. Materiały III Międzynarodowego Kongresu Diagnostyki Technicznej Diagnostyka'2004, Poznań 2004.

[115] Vapnik V.N.: *Statistical Learning Theory*. Wiley, New York, 1998.

[116] Wiśniewski W.: *Diagnostyka techniczna wytwórczych urządzeń energetycznych w elektrowniach*. PWN, Warszawa, 1991.

[117] Wojtusik J.: *Wielomodele diagnostyczne maszyn wirnikowych*. Rozprawa doktorska, Politechnika Śląska, Gliwice, 2006.

[118] Wysogląd B.: *Metody reprezentacji drgań wałów maszyn wirnikowych w diagnostycznych bazach danych*. Politechnika Śląska, Gliwice, 1996. Zeszyty Naukowe – Mechanika z. 126.

[119] Zahn C.T.: *Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters*. IEEE Transactions on Computers, 1/1971, s. 68–86.

[120] Zahn Y., Shen D.: *Design efficient support vector machine for fast classification*. Pattern Recognition, 38/2005, s. 157–161.

[121] Żółtowski B.: *Podstawy diagnostyki maszyn*. Wydawnictwo Uczelniane Akademii Techniczno–Rolniczej, Bydgoszcz, 1996.

[122] Żurada J., Barski M., Jędruch W.: *Sztuczne sieci neuronowe*. PWN, Warszawa, 1996.

Streszczenie

Tworzenie komputerowych systemów wspomagających diagnostykę maszyn wymaga posiadania odpowiedniego źródła reguł, które utworzą bazę wiedzy systemów doradczych. Takim źródłem reguł może być relacja diagnostyczna (model diagnostyczny). Jednym ze sposobów uzyskania modelu diagnostycznego jest otrzymanie go w wyniku odwracania znanego modelu numerycznego rozpatrywanego obiektu. Otrzymany tak model nosi nazwę odwrotnego modelu diagnostycznego.

Metodą umożliwiającą skuteczną identyfikację modeli odwrotnych jest metoda ich identyfikacji na podstawie przykładów otrzymanych z badań symulacyjnych na zaawansowanych modelach numerycznych rozpatrywanego obiektu. Optymalizacja tak identyfikowanych modeli odwrotnych wymaga opracowania skutecznych metod nie tylko wyboru optymalnego zbioru przykładów będących podstawą identyfikacji modelu odwrotnego, ale również wyboru optymalnego zbioru symptomów (cech sygnałów diagnostycznych) uwzględnianych w procesie identyfikacji.

Dla rozwiązania tak określonych zadań w pracy zaproponowano zastosowanie algorytmów ewolucyjnych oraz sformułowano dwie tezy. Pierwsza z nich mówi, że poszukiwanie optymalnego zbioru cech dla identyfikacji odwrotnych modeli diagnostycznych prowadzi do lepszych rezultatów niż poszukiwanie zbioru cech optymalnych. Druga teza stanowi, że nadmierne zwiększanie liczności zbioru cech, na podstawie których zostanie wyznaczony odwrotny model diagnostyczny, prowadzi do spadku jakości wyznaczonego modelu.

Zastosowanie algorytmów ewolucyjnych w zadaniu identyfikacji modeli odwrotnych wymagało rozwiązania wielu zadań cząstkowych: określenia stosownych metod kodowania, przygotowania odpowiednich do przyjętego sposobu kodowania operatorów genetycznych, opracowania funkcji przystosowania umożliwiających otrzymanie poprawnych rozwiązań. Opracowano metodę identyfikacji modeli lokalnych stanowiących pierwszy stopień dwustopniowego modelu odwrotnego bazującą na zaproponowanym algorytmie ewolucyjnym. Opracowano również oprogramowanie umożliwiające zastosowanie opracowanych metod.

Przeprowadzono badania weryfikacyjne wybierając jako obiekt badań turbogenerator energetyczny z turbiną 13K215 mocy 200MW. Otrzymane wyniki potwierdziły prawdziwość postawionych tez.

Summary

To create computer systems aiding machinery diagnosis one requires a suitable source of rules constructing a knowledge base of expert systems is to be built. A diagnostic relation (diagnostic model) may be such a source of the rules. To inverse a known numerical model is one of possible methods of the diagnostic model extraction. The model identified in such the way is called a diagnostic inverse model.

A method enabling us to identify the inverse model efficiently is the method of its identification basing on examples received from simulation research into advanced numerical models of an examined object. Optimization of such identified inverse models requires to elaborate efficient methods of not only optimal choice of an example set being the base of the model identification but also optimal choice of a set of symptoms (diagnostic signal features) considered within the identification process.

In order to solve the task defined in such the way evolutionary algorithms were proposed. Two theses ware stated to be proved. The first one says that searching for an optimal set of features in order to identify diagnostic inverse models, leads to better results than searching for a set of optimal features. The second one states that excessive increasing a size of the set of features, on which the diagnostic inverse model is determined, leads to decrease the quality of a determined model.

The application of evolutionary algorithms to the task of identification of inverse models required many sub-tasks to be solved. Examples are: definition of suitable coding methods, preparation of genetic operators suitable for an accepted coding method, elaboration of fitness function enabling us to obtain correct solutions. The identification method of local models considered as the first stage of two stage inverse model, was elaborated basing on a proposed genetic algorithm. Software enabling the application of the method was also elaborated.

During the verification research of the proposed method a turbo–generator with 200MW of power capacity 13K215 turbine was chosen as the examined object. Obtained results proved the formulated theses.