Politechnika Śląska
Wydział Mechaniczny Technologiczny
Katedra Podstaw Konstrukcji Maszyn

Marcin Bednarski

Metody doskonalenia sieci bayesowskich stosowanych w diagnostycznych systemach doradczych

Gliwice 2006

Recenzenci
Prof. dr hab. inż. Ryszard Michalski, Uniwersytet Warmińsko-Mazurski
Prof. dr hab. Wojciech Moczulski, Politechnika Śląska

Redaktor zeszytów
Wojciech Cholewa

Redaktor techniczny
Marek Wyleżoł

Projekt okładki
Wojciech Cholewa, Marek Wyleżoł

ISBN 83–916957–3–5

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. dra hab. inż. Wojciecha Cholewy. Publiczna obrona rozprawy odbyła się 18 lipca 2006 roku przed Komisją powołaną przez Radę Wydziału Mechanicznego Technologicznego. W opracowaniu zostały uwzględnione uwagi recenzentów rozprawy doktorskiej: prof. dra hab. inż. Ryszarda Michalskiego i prof. dra hab. Wojciecha Moczulskiego.

Składam serdeczne podziękowania Promotorowi rozprawy doktorskiej Panu Profesorowi dr hab. inż. Wojciechowi Cholewie za długoletnią opiekę naukową, inspirację do samodzielnej pracy badawczej, a także wsparcie i pomoc przy realizacji tej pracy.

Chciałbym w tym miejscu szczególnie podziękować Panu Profesorowi dr inż. Wiktorowi Fridowi z Royal Institute of Technology w Sztokholmie, z którego pomocy, głębokiej wiedzy i doświadczenia korzystałem przygotowując część dotyczącą sieci bayesowskiej związanej z diagnozowaniem stanu reaktora w elektrowni jądrowej.

Dziękuję również Koleżankom i Kolegom z Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej za okazaną mi pomoc i życzliwość w czasie wykonywania pracy oraz Monice za to, że jest. Jej dedykuję niniejszą książkę.

Gliwice, październik 2006 *Marcin Bednarski*

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

Spis treści

Wykaz ważniejszych oznaczeń

D	zbiór danych (przykładów),
$Dim(G)$	rozmiar grafu,
$E(A,T;S)$	ważona entropia zbioru przykładów S, ze względu na podział zakresu wartości cechy A za pomocą wartości progowej T,
$Ent(S)$	entropia zbioru przykładów S,
G	acykliczny graf skierowany,
$I(H:X)$	miara ilości informacji,
N	liczba przykładów uczących,
N_{ijk}	liczba przykładów spełniających warunki $X_i = x_i^k$ i $\Pi_i = \pi_i^j$,
$P(A)$	prawdopodobieństwo a priori zdarzenia A
$P(A\|B)$	prawdopodobieństwo warunkowe zdarzenia A jeżeli zaszło zdarzenie B,
$s1$	współczynnik wrażliwości wyznaczany jako pochodna $\frac{d(P(h\|e)(x))}{dx}$,
$s2$	współczynnik wrażliwości reprezentujący zakres zmian $P(h\|e)$,
X	zbiór węzłów dyskretnych,
$x_i^1, \ldots, x_i^{r_i}$	wartości węzła X_i,
α_{ijk}	parametr wykładniczy, często nazywany *hiperparametrem*,
ε	zbiór klas równoważności Markowa,
θ_{ijk}	prawdopodobieństwo $P(X_i = x_i^k \|\Pi_i = \pi_i^j)$,
Θ	zbiór tablic prawdopodobieństw warunkowych,
π_i^j	j-ta kombinacja wartości węzłów nadrzędnych (rodziców) węzła X_i.

Rozdział 1

Wstęp

Nowe systemy komputerowe coraz szerzej wkraczające na obszary zarezerwowane do niedawna wyłącznie dla ludzkiej inteligencji, znajdują zastosowanie praktycznie w każdej dziedzinie, od robotyki przez zarządzenie aż do np. medycyny lub geologii. Dostępność sprzętu komputerowego, opracowanie efektywnych metod programowania oraz bogactwo pomysłów i koncepcji, otwiera nowe potężne możliwości w zakresie złożonego przetwarzania danych, gromadzenia i wykorzystywania wiedzy, symulacji, modelowania i optymalizacji. W badaniach nad sztuczną inteligencją, szeroką dziedzinę stanowi kierunek, koncentrujący się na rozumieniu i odzwierciedlaniu istoty wiedzy oraz sposobie rozumowania eksperta z danej dziedziny. Jego efektem są silnie wyspecjalizowanie systemy doradcze, nazywane również systemami ekspertowymi lub eksperckimi, bazujące na różnego rodzaju reprezentacjach wiedzy [30].

Liczne realizacje systemów doradczych (np. systemy diagnostyczne) charakteryzują się tym, że można je rozpatrywać jako szczególne modele procesów przybliżonego rozumowania specjalistów określonych dziedzin [58]. W ostatnim okresie z coraz większym uznaniem spotykają się modele w postaci sieci bayesowskich, które stanowią skuteczne narzędzie do reprezentacji wiedzy w warunkach niepewności. Znalazły one zastosowanie w wielu dziedzinach, np. w diagnostyce medycznej i technicznej, przy kontroli procesów wytwarzania, dla prognozowania finansowego itp., oraz w różnych gałęziach przemysłu, szczególnie w przemyśle energetycznym, lotniczym, jądrowym, itp. Stosowane są również w szeregu modułach, jak np. moduł Office Assistant oprogramowania Microsoft Office, wykorzystywany w celu doradzania użytkownikowi.

1.1. Geneza pracy

Sieci bayesowskie są szczególnym przypadkiem modeli probabilistycznych. Model w postaci sieci bayesowskiej jest skierowanym, acyklicznym grafem, którego topologia opisuje zależności (oraz ich brak) między zmiennymi modelu reprezentowanymi przez węzły. Każdy węzeł sieci bayesowskiej ma stowarzyszoną z nim tablicę prawdopodobieństw warunkowych określającą zależność pomiędzy węzłami wynikającą z topologii sieci. Sieci bayesowskie, w przypadku diagnostyki technicznej, zwykle wykorzystywane są do tworze-

nia modeli złożonych obiektów technicznych, stosowanych do rozpoznawania stanu technicznego na podstawie niepewnych lub niekompletnych danych. Literatura omawiająca różne zagadnienia związane z sieciami bayesowskimi jest obszerna, np. [36, 54, 55, 64, 88]. Dostępnych jest również wiele programów umożliwiających tworzenie, edycje i wnioskowanie w sieciach bayesowskich [7, 44, 53, 80, 85].

Modele w postaci sieci bayesowskich generalnie identyfikowane są, w większości przypadków, na podstawie wiedzy ekspertów z dziedziny, której dotyczy tworzony model. Inną metodą tworzenia sieci bayesowskiej jest uczenie sieci na podstawie danych w sposób automatyczny. Zadanie to jest złożone obliczeniowo, lecz wykonywalne za pomocą skutecznych algorytmów oraz strategii heurystycznych [24, 32, 48, 49, 82, 93]. W obu tych przypadkach zadanie identyfikacji sieci bayesowskich polega na określeniu struktury oraz prawdopodobieństw warunkowych tzw. parametrów sieci. Identyfikacja taka odbywa się na podstawie danych pochodzących z różnych źródeł, wśród których można wymienić specjalistów, obserwacje diagnostyczne (wyniki eksperymentów czynnych i biernych) oraz eksperymenty numeryczne. Dane te zwykle są niedoskonałe, ponieważ mogą być niekompletne, niepewne, niedokładne i/lub błędne [76]. Dlatego należy liczyć się z pewnymi błędami modelowania spowodowanymi niedoskonałością danych, które w skrajnym przypadku dla systemów diagnostycznych mogą skutkować fałszywymi diagnozami generowanymi przez system.

W przypadku sieci bayesowskich generalnie brak jest opracowań związanych z metodami doskonalenia takich sieci. Natomiast występuje szereg opracowań związanych z analizą wrażliwości sieci bayesowskich [33–35, 55, 60, 63]. Analiza wrażliwości sieci bayesowskich pozwala wyznaczyć wpływ zmiany różnych wielkości związanych z modelem na jakość wnioskowania (wyjście z modelu - prawdopodobieństwo a posteriori hipotez). Takie podejście pozwala na wyznaczenie wielkości, których niedokładność ma duży lub mały wpływ na wyjście z modelu, co w konsekwencji może stanowić podstawę do doskonalenia sieci bayesowskiej.

1.2. Cel rozprawy

Celem niniejszej pracy jest określenie zalet i wad analizy wrażliwości sieci bayesowskiej oraz określenie przydatności takiej analizy do celów doskonalenia sieci bayesowskich. Rozważania te dotyczą diagnostycznych sieci bayesowskich, które mogą być stosowane w diagnostycznych systemach doradczych.

Omawiane zagadnienia będą dotyczyć sieci bayesowskich, zidentyfikowanych zarówno na podstawie danych pochodzących z eksperymentu numerycznego, jak i na podstawie wiedzy ekspertów.

W pierwszym przypadku, sieć bayesowska zidentyfikowana zostanie w procesie uczenia na podstawie danych będących wynikiem symulacji modelu numerycznego turbogeneratora.

W drugim przypadku, badana sieć bayesowska została utworzona w ramach projektu STERPS [46] i dotyczy zagadnienia diagnozowania stanu reaktora w elektrowni jądrowej

oraz oceny ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii.

Zastosowane w badaniach metody analizy wrażliwości zostaną wykorzystane do doskonalenia rozpatrywanych sieci polegającego na optymalizacji struktury oraz optymalizacji parametrów sieci, jako prawdopodobieństw warunkowych, sieci.

1.3. Tezy

W pracy postawiono następujące tezy:

> *Wyniki analizy wrażliwości przeprowadzonej dla węzłów obserwowanych sieci bayesowskiej mogą stanowić podstawę do optymalizacji struktury tej sieci.*

> *Wyniki analizy wrażliwości przeprowadzonej dla parametrów sieci bayesowskiej mogą stanowić podstawę do optymalizacji wartości tych parametrów.*

1.4. Zakres rozprawy

W pracy przedstawiono metodologię doskonalenia diagnostycznych sieci bayesowskich na podstawie dwóch rodzajów analiz wrażliwości. Badania proponowanych metod i ocena ich przydatności, wraz z wnioskami zostały opisane w rozdziałach 5 i 6.
Struktura rozprawy jest następująca:

W rozdziale 2 omówiono zagadnienia związane z sieciami bayesowskimi. Przedstawiono koncepcję sieci bayesowskiej wraz z twierdzeniem Bayesa, na którym została oparta. Scharakteryzowano również przykład diagnostycznej sieci bayesowskiej. Na przykładzie tym zilustrowano różne procesy wnioskowania w takich sieciach. Dodatkowo zwrócono uwagę na właściwość sieci bayesowskiej jaką jest warunkowa niezależność węzłów. Podano jej definicję, jak również zagadnienie to przeanalizowano na przykładzie.

W rozdziale 3 przedstawiono metody identyfikacji sieci bayesowskich. Omówiono proces ustalenia struktury i prawdopodobieństw warunkowych na podstawie wiedzy z dziedziny tworzonego modelu. W rozdziale opisano również proces identyfikacji sieci bayesowskiej na podstawie uczenia sieci z danych. Osobno pokazano proces uczenia struktury sieci i parametrów sieci, przedstawiając najważniejsze algorytmy uczenia. Dokonano przeglądu metod dyskretyzacji danych, ze względu na fakt, że algorytmy uczenia sieci bazują na danych dyskretnych, które w rzeczywistości występują rzadko, w zastosowaniach diagnostycznych. Omówiono również zagadnienie dotyczące oprogramowania do uczenia sieci bayesowskich.

W rozdziale 4 opisano dwa rodzaje analizy wrażliwości sieci bayesowskich: analizę węzłów obserwowanych i analizę parametrów sieci. W pierwszym przypadku zdefiniowana została wrażliwość węzłów obserwowanych. Natomiast, w drugim przypadku omówiono szczegółowo metodę wyznaczania zaproponowanej miary wrażliwości parametru wraz z interpretacją graficzną.

W rozdziale 5 zaprezentowano przykład wykorzystania analizy wrażliwości węzłów obserwowanych do optymalizacji struktury sieci na podstawie kryterium jakości wnioskowania. Przedmiotem badań, w tym przypadku, była sieć bayesowska uzyskana w procesie uczenia sieci na podstawie danych dotyczących przekoszenia panwi łożyskowych oraz przemieszczenia podpór łożyskowych. Przedstawiony proces uczenia sieci został poprzedzony opisem procesu przygotowania danych do procesu uczenia. Wyniki badań w zakresie analizy wrażliwości węzłów obserwowanych oraz optymalizacji struktury zostały podsumowane i sformułowano wnioski.

W rozdziale 6 przedstawiono analizę wrażliwości parametrów sieci bayesowskiej, na podstawie której, jak wykazano, możliwe jest przeprowadzenie procesu doskonalenia tej sieci. Na podstawie wyników przeprowadzonej analizy wykryto wadę zaproponowanej miary wrażliwości. W celu eliminacji tej niedogodności przyjęto dodatkową, nową miarę wrażliwości. Przeprowadzona analiza wrażliwości parametrów sieci bayesowskiej została podsumowana, oraz sformułowano wnioski.

W rozdziale 7 sformułowano ogólne wnioski z przeprowadzonych badań w zakresie metod doskonalenia diagnostycznych sieci bayesowskich, jak również przedstawiono wnioski szczegółowe. Przedstawiono również propozycje kierunków dalszych badań w zakresie doskonalenia sieci bayesowskich.

Rozdział 2

Sieć bayesowska (sieć przekonań)

Sieć bayesowska, również nazywana siecią przekonań, jest intuicyjnym i zbliżonym do ludzkiego postrzegania świata, modelem wybranego fragmentu rzeczywistości. Koncepcja sieci bayesowkiej została zaproponowana przez Pearla [88] i była oraz jest intensywnie rozwijana [5, 17, 36, 54, 55, 64]. Obecnie sieci bayesowskie stały się popularnym narzędziem do reprezentacji wiedzy w warunkach niepewności w wielu dziedzinach zastosowań np. w diagnostyce medycznej [96, 100], w diagnostyce technicznej [13], w systemach doradczych [98], w systemach wspomagających podejmowanie decyzji (computerized decision support systems) [46], przy prognozowaniu finansowym [1], przy kontroli wytwarzania [81], w rozpoznawaniu mowy (language recognition/understanding) [45], w szkoleniu elektronicznym (e-learning) [26] itp.

Istotną zaletą sieci bayesowskich jest przejrzystość oraz czytelność modelu. Z punktu widzenia inżynierii wiedzy, sieć bayesowska może odzwierciedlać strukturę przyczynowo-skutkową, która pozwala na pełniejsze zrozumienie modelowanego problemu zarówno ekspertom jak i użytkownikom systemu. W praktyce takie sieci mogą być projektowane przez ekspertów z danej dziedziny. Graficzne przedstawianie struktury grafu pozwala na łatwe tworzenie lub modyfikację sieci przez eksperta, dzięki czemu można wykorzystywać jego specjalistyczną wiedzę w prosty sposób. Struktury sieci można również odkrywać na podstawie danych, także w sposób automatyczny. Jest to zadanie złożone obliczeniowo, wykonywalne za pomocą skutecznych algorytmów oraz strategii heurystycznych.

2.1. Podstawowe zagadnienia

2.1.1. Twierdzenie Bayesa

Twierdzenie Bayesa[1] jest podstawowym narzędziem teorii prawdopodobieństwa w interpretacji subiektywistycznej tzn. interpretacji według której prawdopodobieństwo nie jest wartością obiektywną, lecz zależy od dostępnych danych (np. prawdopodobieństwo uszkodzenia układu lub systemu technicznego). Prawdopodobieństwo subiektywne nie ma interpretacji częstościowej, ale jest wielkością odzwierciedlającą subiektywne prze-

[1] Thomas Bayes (1702-1761) – brytyjski matematyk

konanie, że pewne zdarzenie może zajść. Prawdopodobieństwo takie obrazuje niepewną wiedzę obserwatora o stanie rzeczy. Twierdzenie Bayesa może być wykorzystane do wyznaczenia takiego prawdopodobieństwa, tzn. znane prawdopodobieństwo a priori hipotezy $P(Hipoteza)$ jest korygowane w świetle dostępnych (ograniczonych) informacji o prawdopodobieństwie zdarzenia $P(Zdarzenie)$ oraz prawdopodobieństwie warunkowym zdarzenia jeżeli hipoteza uznawana jest za prawdziwą $P(Zdarzenie|Hipoteza)$, w wyniku czego uzyskuje się prawdopodobieństwo a posteriori hipotezy (wzór 2.1) [14].

$$P(\textit{Hipoteza} \mid \textit{Zdarzenie}) = \frac{P(\textit{Zdarzenie} \mid \textit{Hipoteza}) \cdot P(\textit{Zdarzenie})}{P(\textit{Zdarzenie})} \tag{2.1}$$

Formalnie twierdzenia Bayesa wyraża następujący wzór:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{P(A)} \tag{2.2}$$

gdzie:

$P(B_i\|A)$	prawdopodobieństwo a posteriori zdarzenia B_i dla danego zdarzenia A,
$P(B_i)$	prawdopodobieństwo a priori zdarzenia B_i (przed uwzględnieniem innych zdarzeń),
$P(A\|B_i)$	prawdopodobieństwo warunkowe zdarzenia A jeżeli zaszło zdarzenie B_i,
$P(A)$	$\sum_j P(B_j) \cdot P(A\|B_j)$ (dla zdarzeń B_j wzajemnie wyłączających się).

2.1.2. Sieć bayesowska

Sieć bayesowska to acykliczny (nie zawierający cykli) graf skierowany, w którym węzły reprezentują zmienne (np. temperaturę jakiegoś źródła, stan pacjenta, cechę obiektu itp.), przyjmujące wartości dyskretne (np.: tak, nie), a skierowane gałęzie zależność pomiędzy tymi zmiennymi. Zależność ta jest określona przez tablicę prawdopodobieństw warunkowych (CPT - conditional probability table), która jest przypisana do każdego węzła w sieci [54].

Formalna definicja sieci bayesowskiej, przedstawiona i omówiona jako definicja 2.1 jest następująca [54]:

Definicja 2.1 *Jeżeli G jest acyklicznym grafem skierowanym (DAG - Directed Acyclic Graph), którego topologia opisuje zależności (oraz ich brak) między zmiennymi modelu reprezentowanymi przez węzły, to przez Π_i rozumiemy zbiór wszystkich kombinacji wartości węzłów nadrzędnych (rodziców) węzła X_i w G. Jeżeli Θ jest zbiorem tablic*

prawdopodobieństw warunkowych, zapisywanych w postaci macierzy θ, takich że zawierają prawdopodobieństwa zdarzeń $P(X_i \mid \Pi_i)$ to parę $B = (G, \Theta)$ nazywamy siecią bayesowską (siecią przekonań).

Sieci bayesowskie są pewną bardziej zwartą reprezentacją rozkładu prawdopodobieństwa zmiennych. Rozkład taki opisuje się równaniem (2.3). Za pomocą tego równania można określić łączny rozkład prawdopodobieństwa $P(X)$ znając jedynie prawdopodobieństwa warunkowe przypisane do każdego węzła.

$$P(X) = \prod_{i=1}^{n} P(X_i|\Pi_i) \tag{2.3}$$

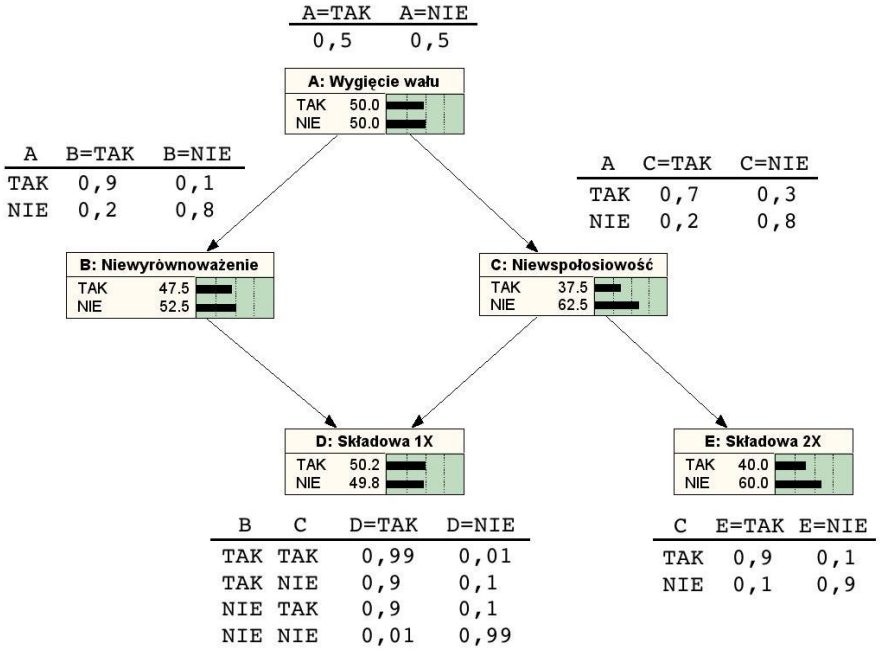
Mając skonstruowaną sieć bayesowską możemy prowadzić różne procesy wnioskowania ogólnie odpowiadające następującemu schematowi:

- część węzłów uznajemy za węzły obserwowane i znamy ich dokładne (pewne) wartości,
- inny zbiór węzłów uznajemy za węzły hipotez i obliczamy prawdopodobieństwo warunkowe hipotezy, że węzeł przyjmuje określoną wartość, względem wartości węzłów obserwowanych.

2.2. Przykład sieci bayesowskiej

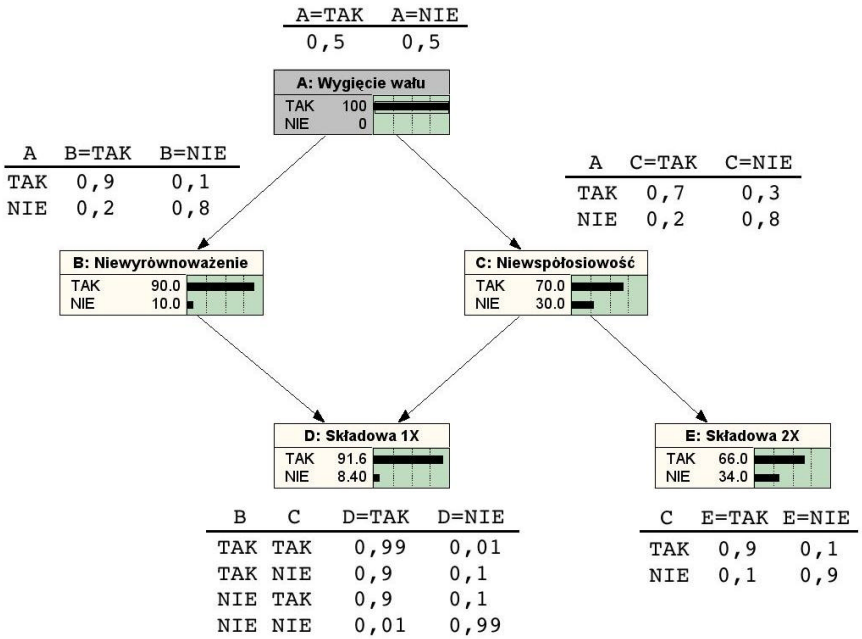
Na rysunku 2.1 przedstawiony został przykład sieci bayesowskiej. Opisuje ona zależności pomiędzy wygięciem wału i niewyrównoważeniem, pomiędzy wygięciem wału a niewspółosiowością, pomiędzy niewyrównoważeniem i niewspółosiowością a wystąpieniem w widmie drgań składowej 1X i pomiędzy niewspółosiowością a wystąpieniem w widmie drgań składowej 2X. Obok każdego węzła przedstawiono tablicę prawdopodobieństw warunkowych określającą każdą z tych zależności. Liczby w każdym węźle przedstawiają odpowiednio rozkłady prawdopodobieństw (określone w procentach). Sieć przedstawiona na rysunku 2.1 reprezentuje sytuację, w której ńie wiemy nic"tzn. nie wiemy czy występuje wygięcie wału, czy występuje niewyrównoważenie, czy występuje niewspółosiowość, czy występuje w widmie drgań składowa 1X i czy występuje w widmie drgań składowa 2X. Przypisanie wartości węzłowi (węzeł koloru szarego) powoduje zmiany rozkładów prawdopodobieństw pozostałych węzłów, których wartości nie zostały ustalone (rysunki 2.2, 2.3, 2.4 i 2.5).
Sieć pokazaną na rysunku 2.1 można rozpatrywać dla dwóch sytuacji:

1. wnioskowanie o symptomach na podstawie znanych uszkodzeń (rysunki 2.2 i 2.3),
2. wnioskowanie o uszkodzeniach na podstawie znanych symptomów (rysunki 2.4 i 2.5.

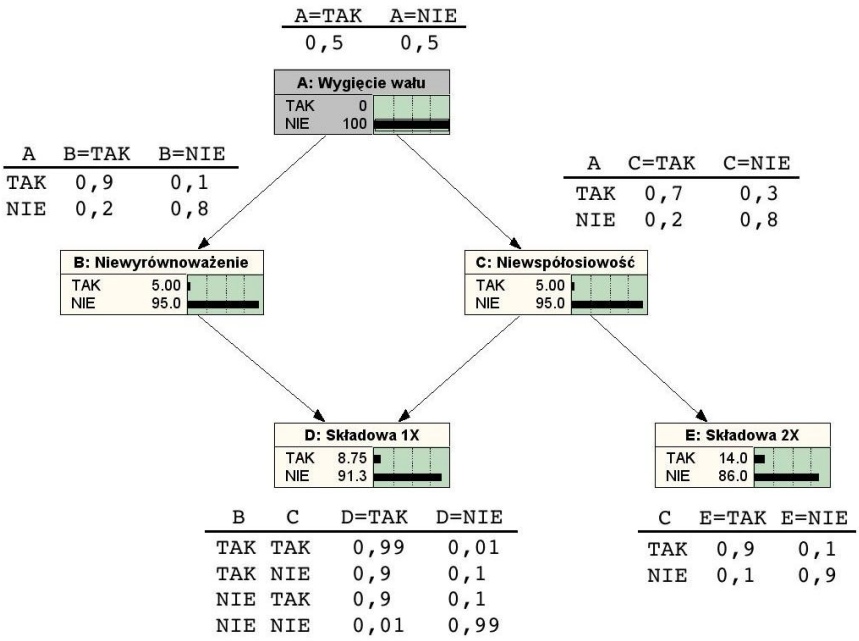


Rys. 2.1. Przykładowa sieć bayesowska

Z porównania rysunku 2.2 i 2.3 (sytuacja 1) wynika, że informacja o tym czy występuje wygięcie wału pozwala wnioskować na temat symptomów, i tak np. jeżeli *Wygięcie wału=TAK* to przekonanie o tym że *Składowa 1X=TAK* (wystąpi składowa 1X) wynosi 91,6%, natomiast w przypadku gdy *Wygięcie wału=NIE* to przekonanie maleje do 8,75%.

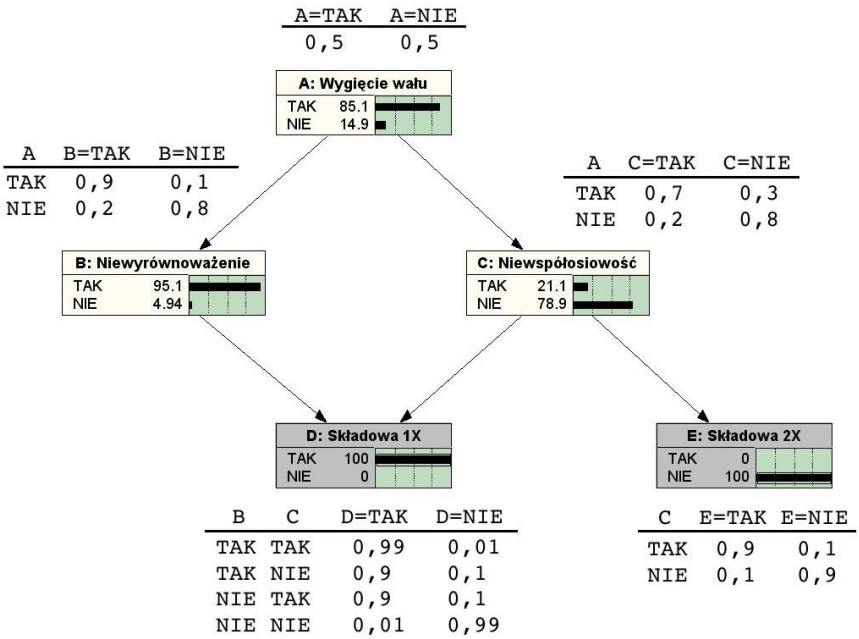


Rys. 2.2. Przykładowa sieć bayesowska dla przypadku gdy "wiemy, że wystąpiło wygięcie wału"

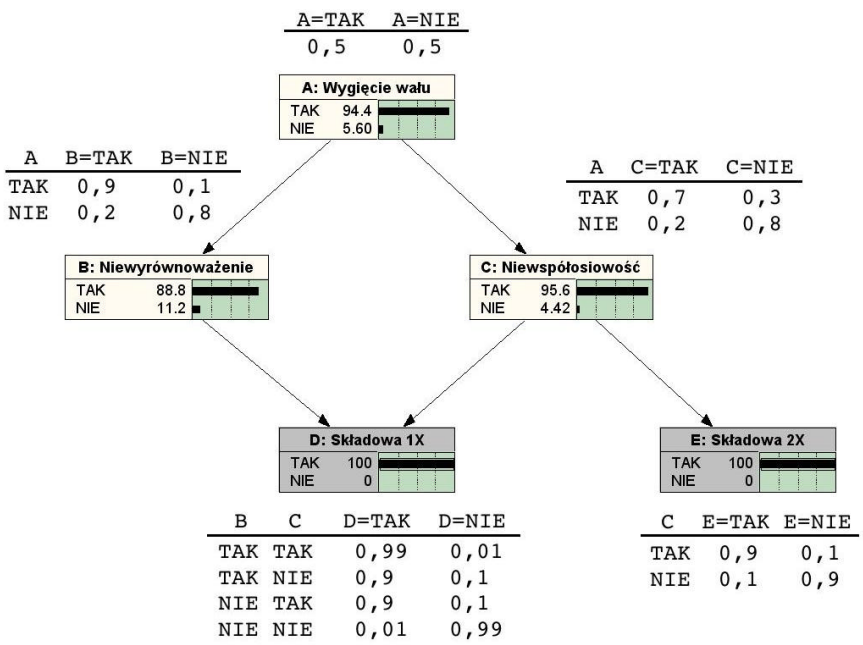


Rys. 2.3. Przykładowa sieć bayesowska dla przypadku gdy "wiemy, że nie wystąpiło wygięcie wału"

Z porównania rysunków 2.4 i 2.5 (sytuacja 2 - wnioskowanie o uszkodzeniach na podstawie znanych symptomów) wynika, że zmiana informacji o tym iż występuje składowa 2X pozwala znacznie zwiększyć przekonanie, że występuje niewspółosiowość (z 21,1% do 95,6%) oraz zmniejszyć przekonanie o tym, że wystąpiło wygięcie wału (z 94,4% do 85,1%) oraz o tym, że wystąpiło niewyrównoważenie (z 95,1% do 88,8%).



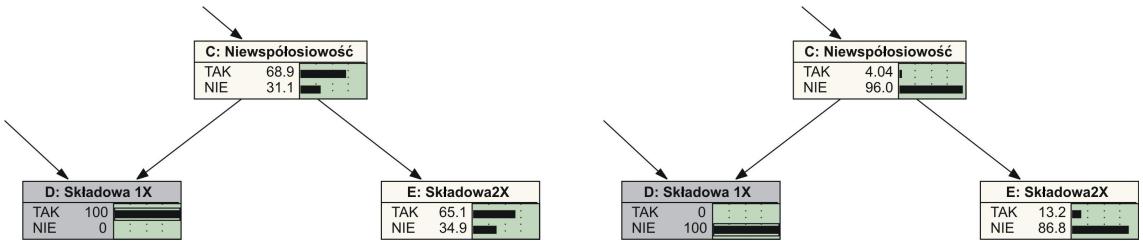
Rys. 2.4. Przykładowa sieć bayesowska dla przypadku gdy "wiemy, że wystąpiła składowa 1X i nie wystąpiła składowa 2X"



Rys. 2.5. Przykładowa sieć bayesowska dla przypadku gdy "wiemy, że wystąpiła składowa 1X i składowa 2X"

2.3. Warunkowa niezależność węzłów

Rozważając sytuację przedstawioną na rysunku 2.6, gdzie nieznana jest wartość węzła *Niewspółosiowość*, możemy zauważyć, że węzły *Składowa 1X* i *Składowa 2X* są warunkowo zależne tzn. zmiana wartości węzła *Składowa 1X* wywołuje zmiany wartości prawdopodobieństw a posteriori węzła *Składowa 2X* i odwrotnie.



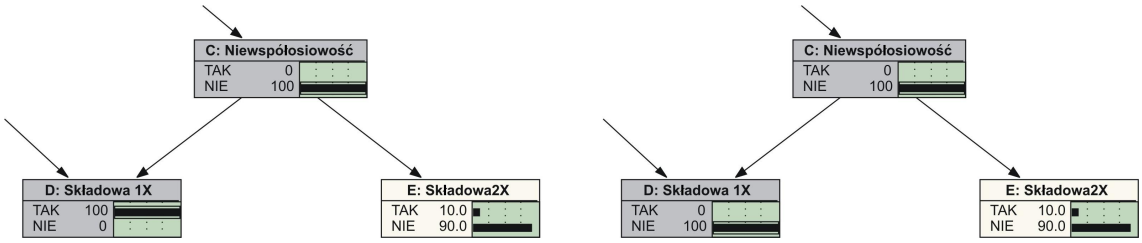
Rys. 2.6. Fragment sieci dla przykładu przedstawionego na rys.2.1. Zmiana wartości węzła *Składowa 1X* wywołuje zmiany wartości prawdopodobieństw a posteriori węzła *Składowa 2X*

Natomiast dla sytuacji gdzie wartość węzła *Niewspółosiowość* jest znana (rys.2.7) węzeł *Składowa 1X* i *Składowa 2X* są warunkowo niezależne tzn. zmiana wartości węzła *Składowa 1X* nie wywołuje zmiany wartości prawdopodobieństw a posteriori węzła *Składowa 2X* i odwrotnie.

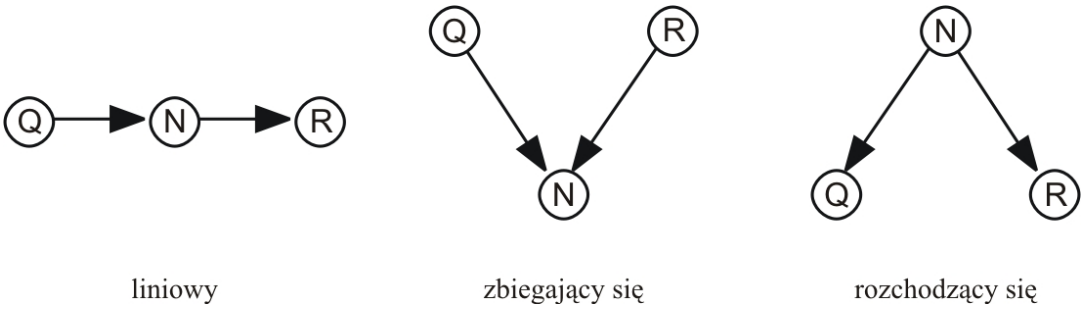
Definicja 2.2 *Dwa (zbiory) węzły Q i R są warunkowo niezależne jeżeli nie istnieje ani jedna ścieżka bezpośrednio łącząca (d-connecting path) Q i R. Ścieżka pomiędzy Q i*

R jest bezpośrednio łączącą (d-connecting) jeśli każdy wewnętrzny węzeł N ścieżki ma jedną z następujących własności:

- *jest liniowy lub rozchodzący się i nie znana jest jego wartość,*
- *jest zbiegający się i wartość jego albo jednego z jego potomków jest znana.*



Rys. 2.7. Fragment sieci dla przykładu przedstawionego na rys.2.1. Zmiana wartości węzła *Składowa 1X* nie wywołuje zmiany wartości prawdopodobieństw a posteriori węzła *Składowa 2X*



Rys. 2.8. Rodzaje węzłów wewnętrznych

Rozdział 3

Definiowanie i identyfikowanie sieci bayesowskich

Modele obiektów oparte na sieciach bayesowskich mogą być identyfikowane na podstawie:

- znanych modeli fizycznych i numerycznych,
- eksperymentów czynnych i biernych,
- opinii ekspertów z danej dziedziny.

Generalnie, w większości przypadków, sieci przekonań identyfikowane są na podstawie wiedzy ekspertów z dziedziny, której dotyczy tworzony model. Alternatywną metodą jest tworzenie modelu na podstawie danych (pozyskiwanie/uczenie sieci), które mogą być pozyskane jako wynik eksperymentów czynnych, biernych i/lub obliczeń symulacyjnych. Większość oprogramowania komputerowego dotyczącego budowania sieci bayesowskich umożliwia, na podstawie danych trenujących, wyznaczenie odpowiednich parametrów sieci (prawdopodobieństw warunkowych). Istnieją również bardziej złożone programy pozwalające wyznaczyć związki pomiędzy zmiennymi reprezentowanymi przez węzły, na podstawie których identyfikowana jest struktura sieci.

Rozdział ten przedstawia opis metody tworzenia sieci bayesowskiej na podstawie wiedzy ekspertów oraz przegląd metod identyfikacji sieci bayesowskiej na podstawie danych. Dodatkowo, dokonano przeglądu metod dyskretyzacji danych, ze względu na fakt, że algorytmy uczenia sieci bazują na danych dyskretnych, które w rzeczywistości występują rzadko, w zastosowaniach diagnostycznych. Na podstawie tego rozdziału dokonano wyboru metody dyskretyzacji i metody uczenia sieci bayesowskiej wykorzystanych w badaniach opisanych w rozdziale 5.

3.1. Ustalenie struktury i prawdopodobieństw warunkowych

Podstawowym założeniem przy budowaniu sieci bayesowskiej jest to, że sieć ta ma wyznaczyć prawdopodobieństwo zdarzeń identyfikujących konkluzje wnioskowania. Pierw-

szym etapem przy definiowaniu sieci bayesowskiej jest więc ustalenie tych zdarzeń. Nazywane są one hipotezami (hypothesis events). Zdarzenia te następnie są grupowane w zbiory wzajemnie wykluczających się zdarzeń, które tworzą węzły hipotez (hypothesis nodes). Kolejnym etapem jest ustalenie jakie informacje związane z hipotezami będą dostępne dla rozważanego modelu - identyfikacja węzłów informacyjnych (information nodes) oraz ich stanów (stwierdzeń). Węzły informacyjne można podzielić na dwie grupy: węzły ukryte (hidden nodes) i węzły obserwowane (observables). Węzły ukryte mogą być wykorzystywane w celu lepszej reprezentacji sieci (np. zapewnienia warunkowej niezależności między niektórymi węzłami, stworzenia struktury umożliwiającej łatwiejsze ustalenie prawdopodobieństw warunkowych itp.). Węzły obserwowane natomiast reprezentują zdarzenia, które mogą być obserwowane.

Kolejnym krokiem po ustaleniu węzłów sieci jest zidentyfikowanie zależności pomiędzy zmiennymi reprezentowanymi przez te węzły (ustalenie gałęzi i ich kierunków). Najczęściej gałęzie te odwzorowują przyczynowo-skutkowe zależności pomiędzy zmiennymi lecz nie jest to regułą, gdyż czasami trudno jest określić co jest przyczyną, a co skutkiem.

Znajomość prawdopodobieństw warunkowych jest najtrudniejszym problem estymacji bayesowskiej. Ogólnie określenie tych prawdopodobieństw może bazować na wiedzy pozyskanej [76]:

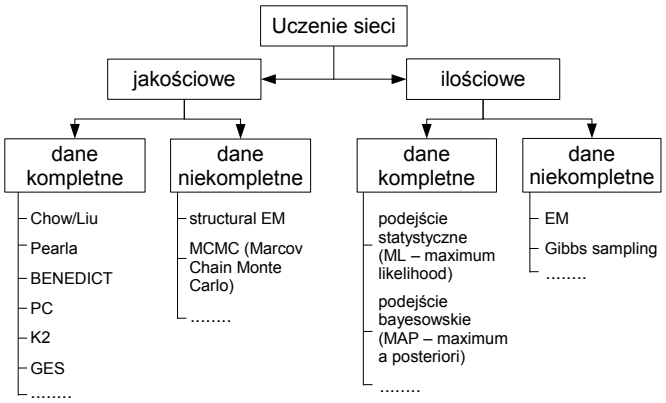
- od specjalistów z zakresu danej dziedziny:
  - bezpośrednio - udział specjalisty w procesie tworzenia sieci bayesowskiej,
  - pośrednio - na podstawie literatury (której autorami są specjaliści) dotyczącej dziedziny modelu;
- z baz danych zawierających:
  - wyniki obserwacji diagnostycznych (wyniki eksperymentów czynnych i biernych [28]),
  - wyniki obliczeń symulacyjnych.

W większości przypadków prawdopodobieństwa warunkowe określane są na podstawie wiedzy i doświadczenia ekspertów z danej dziedziny lub określane są na podstawie wcześniejszych obserwacji identycznych lub podobnych obiektów.

3.2. Identyfikacja sieci bayesowskiej na podstawie danych-uczenie sieci

Uczenie sieci bayesowskich można podzielić na uczenie jakościowe i ilościowe. Jakościowe uczenie dotyczy struktury sieci (gałęzi pomiędzy węzłami), natomiast uczenie ilościowe polega na wyznaczeniu parametrów sieci (tablic prawdopodobieństw warunkowych). Osobnym zagadnieniem, jest uczenie sieci bayesowskich na podstawie niekompletnych danych. W tym celu tworzone są różnego rodzaju algorytmy, z których najbardziej

znany i popularny to algorytm EM [42, 87]. Rysunek 3.1 przedstawia klasyfikacje algorytmów uczenia sieci bayesowskich, które zostały opisane w rozdziale 3.2.3.



Rys. 3.1. Klasyfikacja metod uczenia sieci bayesowskich

W diagnostyce maszyn, zwykle przyjmuje się, że dane są wartościami cech sygnałów i odpowiadającymi im wartościami cech stanu. Cecha definiowana jest jako para [27]:

$$c_i = \langle nam(x_i), val(x_i) \rangle \tag{3.1}$$

gdzie:

$nam(x_i)$ jest nazwą lub indeksem cechy,
$val(x_i)$ jest wartością tej cechy.

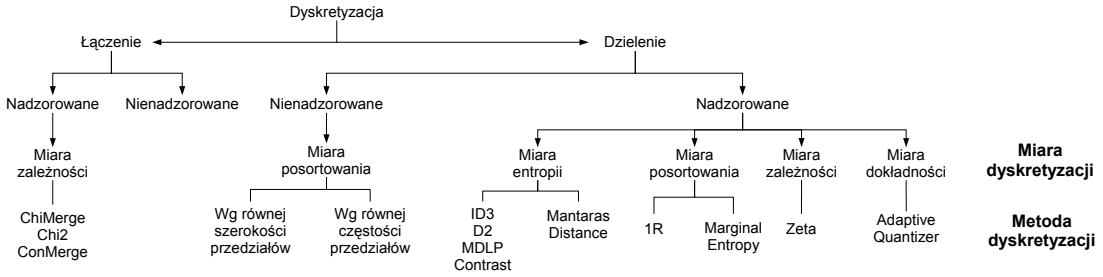
W przypadku sieci bayesowskich, algorytmy uczenia operują na danych dyskretnych. Niestety w większości przypadków wartości cech są ciągłe (wartości będące liczbami rzeczywistymi, np. uzyskane w wyniku pomiaru wielkości fizycznych o charakterze ciągłym). Dlatego też, aby ciągła wartość cechy mogła zostać użyta, musi ona zostać poddana procesowi dyskretyzacji, w wyniku którego otrzymuje się dyskretne wartości cechy.

3.2.1. Dyskretyzacja cech o ciągłych wartościach

Dyskretyzacja jest procesem przekształcania wartości cech o charakterze ciągłym na cechy o wartościach dyskretnych. Celem dyskretyzacji jest dokonanie podziału cechy ciągłej na rozłączne i pokrywające jej zakres przedziały o określonych własnościach. Każdy z takich przedziałów może być wtedy utożsamiany z jedną wartością dyskretną ńowejćechy. Sposób i właściwości, względem których grupowane są wartości zależą od specyfikacji problemu oraz zastosowanego algorytmu.

W literaturze opisywanych jest wiele metod przekształcania cech ciągłych na cechy dyskretne np. [37, 66, 89, 90] itd. Główny podział metod dyskretyzacji jest oparty na kryterium dostępności informacji o klasach: metody z nadzorem (supervised) i bez nadzoru (unsupervised). Kolejnym kryterium klasyfikacji metod dyskretyzacji jest miara dyskretyzacji (discretization measure) np. miara entropii, miara zależności χ^2 itp. Dodatkowo metody dyskretyzacji podzielone są na metody dzielące lub łączące przedziały. Podział

metod dyskretyzacji [66] został przedstawiony na rys.3.2. Poniżej omówiono najważniejsze metody dyskretyzacji, gdzie przyjęto, że dysponujemy zbiorem sklasyfikowanych przykładów reprezentowanych przez wartości cechy ciągłej.



Rys. 3.2. Klasyfikacja metod dyskretyzacji

Podział na przedziały o równej szerokości (equal width)

Podział na k przedziałów o równej szerokości jest najprostszą metodą dyskretyzacji danych. Metoda ta polega na posortowaniu wartości cechy A i podzieleniu tego zbioru na k przedziałów o równej długości, gdzie k jest parametrem podawanym przez użytkownika.

Podział na przedziały o jednakowej częstości występowania (equal frequency)

Metoda ta jest bardzo zbliżona do metody podziału na przedziały o równej szerokości. Zbiór wartości cechy ciągłej dzieli się na k podzbiorów, z których każdy zawiera możliwie jednakową liczbę sąsiednich przykładów ze zbioru uczącego (zbiór należy wcześniej posortować). Wartości maksymalne i minimalne w każdym podzbiorze wyznaczają progi dyskretyzacji.

Metoda ChiMerge [57]

Początkowo każda odmienna wartość cechy jest traktowana jako osobny przedział. Następnie iteracyjnie grupowane są ze sobą sąsiednie przedziały. Statystyka χ^2 używana jest do oceny, czy sąsiednie przedziały można połączyć w jeden. Łączenie przyległych przedziałów jest powtarzane do momentu gdy ich liczba nie osiągnie z góry ustalonej wartości lub gdy wartości statystyk χ^2 wszystkich par przyległych przedziałów są mniejsze niż określona wartość progowa, która jest ustalona przez wybrany poziom istotności.

Metoda Chi2 [67]

Metoda ta jest rozwinięciem metody ChiMerge. Algorytm składa się z dwóch etapów. W pierwszej fazie, podobnie jak w metodzie ChiMerge, wszystkie ciągłe cechy są dyskretyzowane, z tym że zamiast z góry ustalonego progu, którego nie może przekroczyć wartość minimalna statystyki χ^2, próg ten jest sukcesywnie zmniejszany dopóki poziom niekonsystencji (niezgodności) [67] nie przekroczy pewnej wartości δ. Przez niekonsystencję należy rozumieć, że dwa identyczne wzorce zostaną zaklasyfikowane do różnych

klas. W drugiej fazie sprawdzane jest, czy uzyskanego progu nie da się jeszcze zmniejszyć, nie tracąc przy tym zbyt wiele informacji.

Metoda minimalnej długości kodu (MDLP Minimal Description Length Principle) [40]

Jest to jedna z najlepszych i najwydajniejszych metod dyskretyzacji. Metoda ta dzieli rekurencyjnie zbiór wartości rozpatrywanej cechy, wykorzystując do znalezienia progów dyskretyzacji zasadę minimalizacji entropii. Ważona entropia zbioru przykładów S ze względu na podział zbioru wartości cechy A za pomocą wartości progowej T wyraża się następująco:

$$E(A,T;S) = \frac{|S_1|}{|S|}Ent(S_1) + \frac{|S_2|}{|S|}Ent(S_2) \tag{3.2}$$

gdzie:

$$\begin{aligned} S_1 &= a \in S : a < T \\ S_2 &= a \in S : a \geq T \end{aligned} \tag{3.3}$$

Punktem binarnej dyskretyzacji przestrzeni wartości A jest wartość T, dla której wartość ważonej entropii $E(A,T;S)$ jest najmniejsza. O tym, czy daną wartość graniczną można zaakceptować, decyduje warunek stopu, który wykorzystuje zasadę minimalnej długości kodu. Zasada ta opiera się na zagadnieniu znalezienia kosztu komunikacji między nadawcą i odbiorcą. Zakłada się, że nadawca i odbiorca posiadają uporządkowany komplet przykładów. Nadawca dodatkowo zna klasy, do których należą przykłady i musi je przekazać odbiorcy. Nadawca musi więc przesłać do odbiorcy brakującą informację z użyciem minimalnej liczby bitów. Dlatego podział kompletu przykładów na przedziały ma sens wtedy i tylko wtedy gdy koszt i długość wiadomości do wysłania po całym procesie są mniejsze niż przed podziałem. Jeśli podział zwiększa liczbę potrzebnych bitów to dyskretyzacja jest przerywana (spełniony zostaje warunek stopu (3.4)).

$$Gain(A,T_{min};S) < \frac{log_2(N-1)}{N} + \frac{\Delta(A,T_{min};S)}{N} \tag{3.4}$$

gdzie:

$Gain(A,T_{min};S) = Ent(S) - E(A,T_{min};S)$

$\Delta(A,T_{min};S) = log_2(3^k - 2) - [k \cdot Ent(S) - k_1 \cdot Ent(S_1) - k_2 \cdot Ent(S_2)]$

N jest liczbą przykładów w zbiorze S,

k_i jest liczbą klas cechy stanu występujących w S_i.

Metoda D2 [18]

Metoda ta wykorzystuje zasadę minimalizacji entropii, dodatkowo używając kilku warunków stopu. Do tych kryteriów należą m.in.: minimalna liczba przykładów w przedziale, maksymalna liczba przedziałów i minimalna wartość funkcji entropii. Metoda ta

jest przydatna w przypadkach, gdy mamy do czynienia z dużą liczbą przykładów uczących, zawierających wiele cech ciągłych.

Metoda 1R (Holtera) [52]

Dyskretyzacja w tej metodzie polega na posortowaniu wartości dyskretyzowanej cechy i po rozdzielaniu tych wartości na zbiory ze względu na informacje o klasie (metoda z nadzorem). Dodatkowo algorytm nakłada, na każdy podzbiór, warunek (przedział musi zawierać pewną minimalną liczba przykładów), aby nie doprowadzić do przypadku, w którym liczba przedziałów będzie zbliżona do liczby przykładów.

Metoda Mantaras Distance [21]

Metoda ta, rekurencyjnie dzieli zbiór wartości cechy na przedziały ze względu na minimalną odległość Mantarasa [71]

$$Dist(S_1, S_2) = \frac{I(S_1|S_2) + I(S_2|S_1)}{I(S_1 \cap S_2)} \tag{3.5}$$

gdzie: $I(S_1|S_2)$, $I(S_2|S_1)$, $I(S_1 \cap S_2)$ są miarą informacji Shannona (entropii warunkowej). Warunkiem stopu w tej metodzie jest kryterium minimalnej długości kodu (MDLP).

Metoda Zeta [51]

Metoda ta bazuje na mierze Z (zeta) (3.6), która definiuje zależność pomiędzy cechami dyskretyzowanymi a informacją o klasie. Metoda ta dzieli rekurencyjnie dziedzinę wartości rozpatrywanej cechy na dwa przedziały wybierając punkt podziału dla którego wartość Z jest największa. Warunek stopu jest definiowany jako maksymalna liczba przedziałów.

$$Z = \sum_{i=1}^{k} n_{f(i),i} \tag{3.6}$$

gdzie:

- k liczba wcześniej określonych przedziałów (domyślnie 2),
- $f(i)$ indeks klasy, zawierający najwyższy obliczony przypadek w i-tym przedziale,
- $n_{(i),i}$ liczba przypadków w i-tym przedziale z indeksem klasy $f(i)$.

Metoda Adaptative Quantizer [22]

Metoda ta jest połączeniem metod dyskretyzacji z nadzorem i bez nadzoru. Na początku dziedzina ciągłej cechy jest dzielona na dwa przedziały o równej szerokości lub jednakowej częstości, a następnie dokonywane są testy oceniające jakość klasyfikacji dla takiego podziału (wybór klasyfikatora zależy od użytkownika). Podzbiór, dla którego jakość klasyfikacji jest mniejsza dzielony jest ponownie. Procedura ta jest powtarzana tak długo dopóki istnieje zadowalająca (większa od założonego progu) poprawa jakości klasyfikacji.

3.2.2. Uczenie ilościowe

Uczenie parametrów sieci bayesowskich (prawdopodobieństw warunkowych) sprowadza się głównie do zliczania liczby przypadków dla różnych warunków (kombinacji stanów rozpatrywanego węzła i jego rodziców). W przypadku zbiorów z niewielką liczbą rekordów, niejednokrotnie dana kombinacja wartości w tabeli prawdopodobieństw warunkowych jest reprezentowana przez zaledwie kilka przykładów (bądź też nawet przez zerową liczbę przykładów), które nie pozwalają na wyznaczenie wiarygodnych parametrów. W takich sytuacjach często stosuje się rozkłady jednostajne [48].

Niech $D = d^1, \ldots, d^N$ oznacza zbiór danych kompletnych (przykładów) zawierających r_i możliwych wartości $x_i^1, \ldots, x_i^{r_i}$, dla $i = 1, \ldots, n$, zmiennych dyskretnych $X = X_1, X_2, \ldots, X_n$. Niech θ_{ijk} oznacza prawdopodobieństwo $P(X_i = x_i^k | \Pi_i = \pi_i^j)$, gdzie π_i^j jest j-tą kombinacją wartości węzłów nadrzędnych (rodziców) węzła X_i. Podobnie, przez $\theta_{ij} = \theta_{ijk} | 1 \leq k \leq r_i$ oznaczono zbiór parametrów opisujących rozkład $P(X_i | \pi_i^j)$.

W klasycznym podejściu statystycznym, parametry (prawdopodobieństwa) uważane są jako własności badanych zjawisk. Zakłada się, że są one objektywnymi stałymi, które w przypadku sieci bayesowskich mogą być wyznaczone na podstawie zbioru przykładów trenujących z wykorzystaniem oceny maksymalnej wiarygodności (ML - maximum likelihood). Logarytm wiarygodności $lnP(D|\Theta)$ może być zdekomponowany zgodnie ze strukturą grafu G i równaniem opisującym łączny rozkład prawdopodobieństwa (równanie 2.3):

$$lnP(D|\Theta) = \sum_{i,j,k} N_{ijk} ln\theta_{ijk} \tag{3.7}$$

gdzie: N_{ijk} reprezentuje liczbę przykładów spełniających warunki $X_i = x_i^k$ i $\Pi_i = \pi_i^j$. Wartość funkcji 3.7 jest maksymalna dla częstości [48]:

$$\hat{\theta}_{ijk} = \frac{N_{ijk}}{N_{ij}} \tag{3.8}$$

gdzie N_{ij} jest liczbą przykładów spełniających warunek $\Pi_i = \pi_i^j$, stąd

$$N_{ij} = \sum_{k=1}^{r_i} N_{ijk} \tag{3.9}$$

W przypadku podejścia bayesowskiego, parametry reprezentują subiektywny stopień przekonania. W takim przypadku zakłada się, że pierwotne przekonanie (prior belief) o θ bazujące np. na wiedzy lub informacjach historycznych jest reprezentowane przez pierwotny rozkład (prior distribution) $P(\theta)$. W momencie gdy dostępne są nowe dane, przekonanie θ jest modyfikowane zgodnie z twierdzeniem Bayesa:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{3.10}$$

Popularnym sposobem reprezentacji pierwotnych przekonań jest rozkład Dirichleta, definiowany jako:

$$Dir(\theta|\alpha_{ij1},\ldots,\alpha_{ijr_i}) = \frac{\Gamma(\alpha_{ij})}{\prod_{k=1}^{r_i}\Gamma(\alpha_{ijk})}\prod_{k=1}^{r_i}\theta_{ijk}^{\alpha_{ijk}-1} \tag{3.11}$$

gdzie $\alpha_{ij} = \sum_{k=1}^{r_i}\alpha_{ijk}$ a $\Gamma(\cdot)$ jest funkcją Gamma. Parametr wykładniczy α_{ijk} jest często nazywany *hiperparametrem*, który interpretowany jest jako liczba przykładów wcześniej obserwowanych przez ęksperta" dla których $X_i = x_i^k$ i $\Pi_i = \pi_i^j$. Z tego powodu parametry α nazywane są również równoważnym rozmiarem przykładów (equivalent sample size). Dla danej struktury G, kompletnych danych D i zbioru hiperparametrów α_{ijk}, wartość oczekiwana parametru sieci w odniesieniu do maksymalnej wartości a posteriori (MAP - maximum a posteriori) $P(\theta|D,G,\alpha_{ij})$ może być wyznaczona na podstawie:

$$\hat{\theta_{ijk}} = \frac{\alpha_{ijk}+N_{ijk}}{\alpha_{ij}+N_{ij}} \tag{3.12}$$

Ustalenie wszystkich parametrów α_{ijk} dla wszystkich możliwych kombinacji (x_i^k, π_i^j) jest trudne i w większości przypadków niepotrzebne. Dlatego też, duża część algorytmów uczenia parametrów sieci bayesowskiej wykorzystuje nieinformacyjne ustalanie parametrów np. w [32] zaproponowano rozkład jednostajny z $\alpha_{ijk} = 1$ natomiast w [15] zaproponowano przyjęcie $\alpha_{ijk} = \alpha/(r_i \cdot q_i)$, gdzie r_i jest liczbą możliwych wartości X_i, q_i jest liczbą możliwych kombinacji Π_i, a α jest równoważnym rozmiarem przykładów.

3.2.3. Uczenie jakościowe

W literaturze dotyczącej sieci bayesowskich [48, 49, 55, 82] możemy spotkać dwa podejścia do problemu uczenia: metody oparte na testach warunkowej niezależności oraz metody oparte na miarach oceny sieci bayesowskiej i procedurze przeszukiwania. Istnieje również szereg metod hybrydowych łączących oba podejścia.

Badania jakościowych zależności i niezależności pomiędzy zmiennymi, reprezentowanymi przez węzły sieci, według algorytmów opartych na testach warunkowej niezależności pozwalają na identyfikację struktury sieci bayesowskiej, która w jak najlepszym stopniu reprezentuje te zależności. Opis algorytmów bazujących na tym podejściu można znaleźć miedzy innymi w [16, 23].

Z zastosowaniem algorytmów opartych na mierze oceny sieci bayesowskiej poszukuje się grafu, spośród wszystkich możliwych grafów, którego jakość wyznaczona na podstawie miary oceny sieci bayesowskiej jest największa. Miara oceny sieci bayesowskiej jest najczęściej definiowana jako miara dopasowania struktury grafu do rzeczywistych danych. Miara ta może bazować na różnych podejściach, takich jak: entropia [50], podejście bayesowskie [32, 49] lub minimalna długość kodu [12, 62]. Najbardziej popularne miary oceny sieci bayesowskiej (metryki) to:

miara AIC (Akaike Information Criterion) [4],

$$Q_{AIC}(G,D) = Ent(G,D) + Dim(G) \tag{3.13}$$

miara BIC (Bayesian Information Criterion) [92],

$$Q_{BIC}(G,D) = Ent(G,D) + \frac{1}{2}Dim(G)lnN \tag{3.14}$$

miara bayesowska (Bayesian metric) [49].

$$Q_{Bayesian}(G,D) = P(G)\prod_{i=1}^{n}\prod_{j=1}^{q_i}\frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij}+N_{ij})}\prod_{k=1}^{r_i}\frac{\Gamma(\alpha_{ijk}+N_{ijk})}{\Gamma(\alpha_{ijk})} \tag{3.15}$$

gdzie:

$Ent(G,D)$ jest entropią pomiędzy strukturą grafu i danymi,

$$Ent(G,D) = -N\prod_{i=1}^{n}\prod_{j=1}^{q_i}\prod_{k=1}^{r_i}\frac{N_{ijk}}{N}ln\frac{N_{ijk}}{N_{ij}} \tag{3.16}$$

$Dim(G)$ jest rozmiarem grafu (liczbą parametrów),

$$Dim(G) = \sum_{i=1}^{n}(r_i - 1)q_i \tag{3.17}$$

$P(G)$ jest prawdopodobieństwem a priori grafu reprezentującym informacje o strukturze sieci, w większości algorytmów prawdopodobieństwo to jest pomijane, gdyż wszystkie oceniane grafy uznaje się jako równie prawdopodobne,

N jest liczbą przykładów uczących.

Badanie pola możliwych rozwiązań odbywa się za pomocą metod przeszukiwania heurystycznego opisywanych w literaturze dotyczącej sztucznej inteligencji (np. [75]). Każda metoda uczenia jakościowego jest więc opisywana poprzez miarę oceny sieci bayesowskiej i procedurę przeszukiwania.

Metody uczenia struktury sieci bayesowskich są od kilku lat intensywnie rozwijane przez ciągłe poszukiwania bardziej wydajnych i dokładniejszych metod [93]. Główne i najbardziej znane algorytmy można przedstawić następująco:.

Algorytm Chow/Liu (Maximum weight spanning tree) [31]

Jest to pierwszy i zarazem najprostszy algorytm. Tworzy sieć o strukturze drzewa skierowanego (kierunki dobierane są w sposób dowolny). Algorytm opiera się na minimalizacji funkcji entropii. Jako entropię wykorzystano funkcję entropii Kullbacka-Leiblera.

Algorytm Pearla [88]

Wynikiem zastosowania algorytmu jest sieć o strukturze polidrzewa (sieć bayesowska o nieskierowanej strukturze, mającej postać drzewa). Według algorytmu, sieć jest identyfikowana podobnie jak w przypadku algorytmu Chow/Liu, jednak różni się tym, że mogą w niej występować węzły zbiegające się.

Algorytm BENEDICT (BElief NEtworks DIscovery using Cut-set Techniques) [2]

Algorytm ten łączy w sobie zarówno elementy klasy algorytmów opartych na entropii, jak i elementy klasy algorytmów opartych na testach niezależności. Główną ideą algorytmu jest wyznaczanie niezgodności pomiędzy warunkowymi niezależnościami reprezentowanymi przez rozpatrywaną siecią bayesowską a tymi, które występują w danych. Do wyznaczania niezgodności wykorzystywana jest miara entropii Kullbacka-Leiblera. Miara globalnej niezgodności jest następnie wykorzystywana w heurystycznej metodzie przeszukiwania przestrzeni możliwych rozwiązań.

Algorytm PC [97]

Algorytm ten oparty jest na testach warunkowej niezależności pomiędzy węzłami. Sieć bayesowska jest przebudowywana, na podstawie danych, przez lokalne badanie fragmentów acyklicznego grafu skierowanego. W pierwszej fazie tworzony jest graf nieskierowany przy założeniu, że jeśli w strukturze sieci węzły są niepołączone, to muszą być warunkowo niezależne. W kolejnych fazach gałęzie są orientowane na podstawie warunkowej niezależności węzłów sieci. Algorytm ten posiada wiele odmian i udoskonaleń np. algorytm BN-PC-B [24]

Algorytm K2 [32]

Algorytm ten należy do grupy algorytmów heurystycznych wykorzystujących miarę bayesowską jako funkcję oceniającą. Funkcja ta określa stopień dopasowania konfiguracji danego węzła X_i do rzeczywistego rozkładu, przy czym pod uwagę bierze jedynie gałęzie skierowane od węzłów ze zbioru rodziców węzła X_i do węzła X_i. Do każdego węzła dodawani są nowi rodzice, aż przestanie to przynosić poprawę wartości funkcji dopasowania lub aż osiągnięta zostanie maksymalna liczba rodziców określona przez użytkownika.

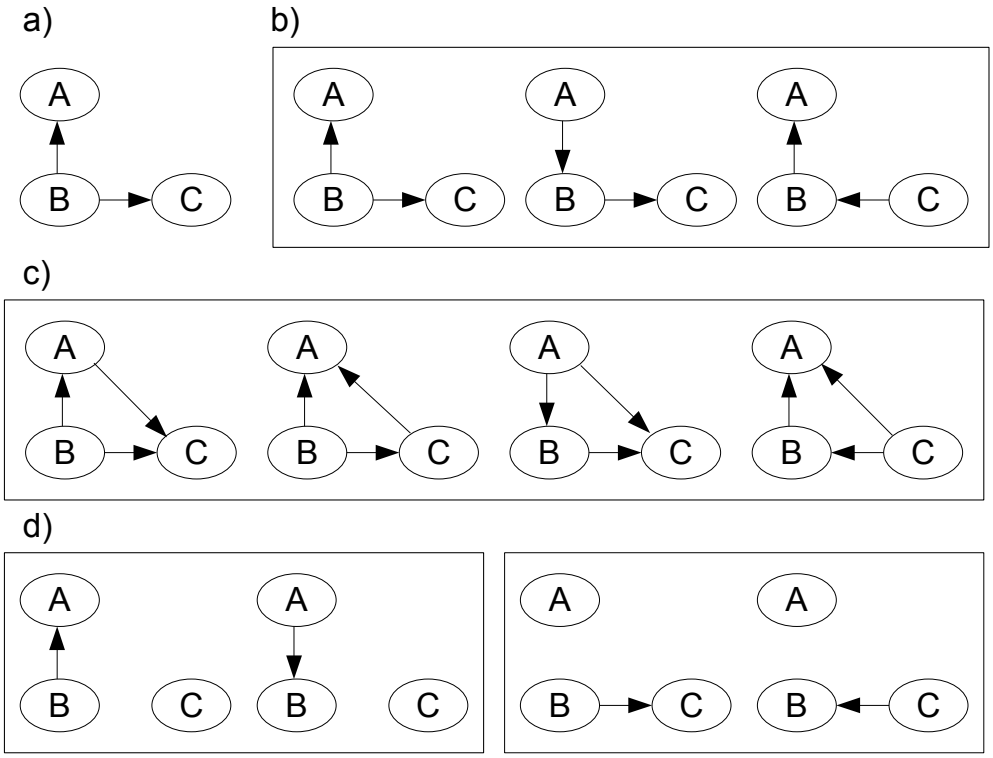
Algorytm GES (Greedy Equivalence Search) [25]

Algorytm ten składa się z dwóch etapów przeszukiwania przestrzeni klas równoważności Markowa (Markov equivalence classes) acykliczych grafów skierowanych.

Definicja 3.1 *Dwa acykliczne grafy skierowane tworzą klasę równoważności ($\equiv$), jeśli reprezentują ten sam rozkład prawdopodobieństwa łącznego. Zbiór klas równoważności Markowa (ε) jest definiowany jako $\varepsilon = A/\equiv$, gdzie A oznacza zbiór wszystkich acyklicznych grafów skierowanych.*

Przeszukiwanie zachłanne (greedy search) rozpoczyna się od grafu bez żadnych gałęzi. W pierwszym etapie powtarzany jest krok zastępowania klasy równoważności ε przez element ze zbioru $\varepsilon^+(\varepsilon)$ (zbiór klas równoważności Markowa utworzony przez dodanie jednej gałęzi pomiędzy węzłami w klasie ε) dla którego wartość miary oceny sieci bayesowskiej jest największa. Dodawanie nowych gałęzi następuje, aż do momentu gdy

nowa klasa równoważności nie zwiększa wartości miary oceny sieci bayesowskiej. W drugim etapie, podobnie jak w pierwszym, zastępowane są kolejne klasy ε przez element ze zbioru $\varepsilon^-(\varepsilon)$ (zbiór klas równoważności Markowa utworzony przez usunięcie jednej gałęzi pomiędzy węzłami w klasie ε), dopóki wartość miary oceny sieci bayesowskiej się zwiększa. Na rysunku 3.3a,b przedstawiono przykład acyklicznego grafu skierowanego G oraz zbiór równoważności Markowa $\varepsilon(G)$. Dodatkowo na rysunku 3.3c,d przedstawiono jednoelementowy zbiór $\varepsilon^+(\varepsilon)$ i dwuelementowy zbiór $\varepsilon^-(\varepsilon)$.



Rys. 3.3. Przykład: a) acyklicznego grafu skierowanego G, b) zbioru $\varepsilon = \varepsilon(G)$, c) zbioru $\varepsilon^+(\varepsilon)$ i d) zbioru $\varepsilon^-(\varepsilon)$

3.2.4. Oprogramowanie

Istnieje wiele programów umożliwiających uczenie sieci. W większości przypadków programy takie umożliwiają tylko wyznaczanie prawdopodobieństw warunkowych (np. Netica [85]), do struktury wcześniej ustalonej (np. wg rozdz.3.1). Dostępne są również programy pozwalające na tworzenie, na podstawie danych pełnej sieci przekonań, tzn. struktury i parametrów (np. Belief Network PowerConstructor [10], Bayes Net Toolbox for Matlab [7]).

Rozdział 4

Analiza wrażliwości sieci bayesowskich

Analiza wrażliwości dotyczy badania wpływu zmian pewnych parametrów modelu na jego zachowanie. W przypadku sieci bayesowskich, analiza wrażliwości dotyczy badania wpływu zmiany pewnych wielkości związanych z modelem (wartości węzłów obserwowanych w przypadku analizy wrażliwości węzłów obserwowanych lub wartości prawdopodobieństw warunkowych w analizie wrażliwości parametrów) na zmiany prawdopodobieństw a posteriori hipotez.

4.1. Analiza wrażliwości węzłów obserwowanych

Analiza wrażliwości węzłów obserwowanych pozwala na odróżnienie ważnych węzłów obserwowanych od takich, które niewiele wnoszą do wyniku działania sieci bayesowskiej. Analizę taką można wykonywać po to by uzyskać wgląd w użyteczność poszczególnych węzłów obserwowanych lub w celu wskazania węzłów, które bez straty jakości sieci bayesowskiej mogą być pominięte. Z pewnych względów, konieczna jest jednak ostrożność przy pomijaniu takich węzłów. Usunięcie jednego z węzłów, może przy pewnych warunkach (niekompletne, niepewne, błędne dane), prowadzić do sytuacji, w której niemożliwe będzie prowadzenie wiarygodnego procesu wnioskowania. Dodatkowo, na podstawie analizy wrażliwości węzłów obserwowanych, można wskazać węzły kluczowe, których nigdy nie wolno pomijać. Analizę wrażliwości węzłów obserwowanych można wykonać korzystając z programów związanych z sieciami bayesowskimi (np. Netica [85]).

W analizie wrażliwości węzłów obserwowanych, w pierwszej kolejności określana jest warunkowa niezależności węzłów obserwowanych i węzłów hipotez, na podstawie której wyznaczane są węzły niewrażliwe. W kolejnym kroku bada się stopień zależności pomiędzy pozostałymi węzłami obserwowanymi a węzłami hipotez, wykorzystując miarę entropii lub współczynnik wzajemnej informacji [94, 101]. Należy podkreślić, że analiza wrażliwości węzłów obserwowanych zależy od kontekstu tzn. przypadku wnioskowania - ustalonych wartości węzłów obserwowanych.

4.1.1. Definicja

Formalna definicja wrażliwości węzłów obserwowanych jest następująca (na podstawie [55]:

Definicja 4.1 *Niech e oznacza zbiór informacji o zdarzeniach obserwowanych, h hipotezę, a O będzie kontekstem (zbiorem ustalonych węzłów obserwowanych). Mówimy, że zbiór zdarzeń obserwowanych $e' \subseteq e$ jest* **wystarczający** *dla danego kontekstu O, jeżeli $P(h|e')$ jest prawie równe $P(h|e)$. Wówczas możemy również mówić, że $e \backslash e'$ jest zbiorem* **nadmiarowym**. *Termin "prawie równy"można zdefiniować poprzez ustalenie progu κ_1 i warunku, że $\left|\frac{P(h|e')}{P(h|e)} - 1\right| < \kappa_1$.*

Na podstawie powyższej definicji można wprowadzić dodatkowe określenia:

Minimalny zbiór wystarczający to taki zbiór e', który jest zbiorem wystarczającym i nie istnieje żaden podzbiór właściwy zbioru e', który jest zbiorem wystarczającym, dla danego kontekstu O.

Zbiór kluczowy to taki zbiór e', który jest podzbiorem, każdego zbioru wystarczającego, dla danego kontekstu O.

Zbiór znaczący to taki zbiór e', dla którego $\left|\frac{P(h|e \backslash e')}{P(h|e)} - 1\right| > \kappa_2$, dla danego kontekstu O, gdzie κ_2 jest pewnym przyjętym progiem.

4.2. Analiza wrażliwości parametrów (prawdopodobieństw warunkowych)

Analiza wrażliwości parametrów (prawdopodobieństw warunkowych) jest szczególnie użyteczna w procesie weryfikacji sieci bayesowskiej, ponieważ jak wiadomo, określenie prawdopodobieństw warunkowych składających się na tablice prawdopodobieństw warunkowych (CPT - conditional probability tables), jest kluczowym zagadnieniem w procesie definiowania sieci bayesowskich. Określenie prawdopodobieństw warunkowych najczęściej bazuje na wiedzy i intuicji ekspertów z danej dziedziny. Niestety w takim przypadku, jakość oraz dokładność uzyskanych prawdopodobieństw jest nieznana. W wielu przypadkach duża dokładność prawdopodobieństw warunkowych nie jest wymagana, lecz z drugiej strony istnieją takie prawdopodobieństwa warunkowe, które mają duży wpływ na wyniki wnioskowania w sieciach bayesowskich. Prawdopodobieństwa te nazywane są słabymi punktami sieci bayesowskiey (weak points of the network).

Zagadnienie analizy wrażliwości parametrów sieci bayesowskiej przedstawiane jest w literaturze dosyć rzadko. Najbardziej znaczące pozycje to [33–35, 55, 60, 63]. Analiza wrażliwości parametrów sieci bayesowskiej pozwala na wyznaczenie listy parametrów wrażliwych, która może stanowić, dla ekspertów z danej dziedziny, wytyczną do określania bardziej dokładnie niektórych prawdopodobieństw warunkowych w procesie optymalizacji (dostrajania) sieci bayesowskiej.

4.2.1. Opis metody

Rozpatrywana analiza wrażliwości parametrów sieci bayesowskiej jest metodą analizy wrażliwości pierwszego rzędu tzn. badany jest wpływ zmian, na wartości prawdopodobieństw hipotez, wyłącznie jednego parametru, przy czym nie rozważane są kombinacje parametrów [33,55]. Przedstawiona poniżej metoda analizy wrażliwości parametrów przebiega dwuetapowo:

etap 1 - wyznaczenie zbioru węzłów, których zmiana parametrów wpływa na wartości prawdopodobieństw a'posteriori w założonych węzłach hipotez i dla ustalonych wartości węzłów obserwowanych tzw. (*zbioru węzłów wrażliwych*),

etap 2 - identyfikacja relacji pomiędzy rozważanymi parametrami (prawdopodobieństwami warunkowymi) i prawdopodobieństwami każdej hipotezy węzła hipotez.

W przypadku pełnej analizy wrażliwości pierwszego rzędu sieci bayesowskiej, należy wyznaczyć wrażliwość dla każdej pary (parametru i hipotezy). Każda taka para składa się z prawdopodobieństwa warunkowego i prawdopodobieństwa a posteriori hipotezy. Obliczenia prowadzi się oddzielnie dla każdej takiej pary i dla każdego kontekstu (przypadku wnioskowania - ustalonych wartości węzłów obserwowanych). Niestety takie podejście jest bardzo czasochłonne nawet dla małych sieci. Zakres obliczeń może być jednak znacząco zredukowany przez pominięcie w obliczeniach węzłów niezależnych, tzn. takich, których prawdopodobieństwa warunkowe nie mają wpływu na wartości prawdopodobieństw hipotez dla danego kontekstu. Węzły takie mogą być wyznaczone na podstawie badania struktury z wykorzystaniem warunkowej niezależności węzłów, bez potrzeby przeprowadzania pełnych obliczeń w sieci bayesowskiej. Węzły, które muszą być wzięte pod uwagę i nie mogą zostać pominięte w dalszych obliczeniach tworzą *zbioru węzłów wrażliwych*. Formalna definicja *zbióru węzłów wrażliwych* jest następująca (na podstawie [33]):

Definicja 4.2 *Niech $B = (G, \Theta)$ będzie siecią bayesowską, gdzie G jest acyklicznym grafem skierowanym a Θ jest zbiorem tablic prawdopodobieństw warunkowych. Niech $V_r \in V(G)$ będzie rozpatrywanym węzłem hipotez a $O \subseteq V(G)$ będzie kontekstem (zbiorem ustalonych węzłów obserwowanych). Niech G^* będzie grafem skierowanym utworzonym z G przez dodanie dodatkowych węzłów nadrzędnych (rodziców) X_i do każdego węzła $V_i \in V(G)$. Wtedy* **zbiorem węzłów wrażliwych** *$Sen(V_r, O)$ dla węzła V_r i kontekstu O jest zbiór wszystkich węzłów V_i dla których X_i (w G^*) jest warunkowo zależny od V_r dla danego kontekstu O.*

Dodatkowy węzeł nadrzędny X_i, może być utożsamiany z niedokładnościami oszacowań prawdopodobieństw warunkowych węzła V_i. Jeśli więc V_r (rozpatrywany węzeł hipotez) jest warunkowo zależny od węzła reprezentującego niedokładności dla danego kontekstu, to zmiana prawdopodobieństwa warunkowego węzła V_i może wpływać na prawdopodobieństwa a posteriori węzła V_r. Dlatego też węzeł V_i należy do *zbioru węzłów wrażliwych*.

W celu identyfikacji zależności pomiędzy parametrem x a prawdopodobieństwem a posteriori hipotezy $P(H = h|e)$ należy posłużyć się następującym równaniem [55] wynikającym z własności sieci bayesowskich:

$$P(H = h|e)(x) = \frac{P(h,e)}{P(e)} = \frac{c_0 + c_1 x}{d_0 + d_1 x} \tag{4.1}$$

gdzie:

$P(H = h\|e)$	prawdopodobieństw a posteriori hipotezy $H = h$,
x	parametr (prawdopodobieństwo warunkowe),
e	kontekst.

Identyfikacja równania (4.1) polega na wyznaczeniu nieznanych współczynników c_0, c_1, d_0, d_1. Przy wyznaczaniu tych współczynników należy zwrócić uwagę na fakt, że parametr x przypisany do węzła A jest związany z innymi parametrami tego węzła. Zmieniając więc wartość parametru x przestaje być spełniona zależność (4.2) dotycząca prawdopodobieństw warunkowych:

$$\sum_{i=1}^{i=N} P(x_i|\pi) = 1 \tag{4.2}$$

gdzie N jest liczbą stanów węzła A, a π konfiguracją stanów węzłów nadrzędnych (rodziców) węzła A. Zmieniając więc parametr x (w celu wyznaczenia współczynników c_0, c_1, d_0, d_1), aby spełnić warunek (4.2), należy zmodyfikować (poddać normalizacji) parametry węzła A zgodnie z:

$$P^*(a_i|\pi)(x^*) = P(a_i|\pi)(x^*)\frac{1}{\sum\limits_i P(a_i|\pi)(x^*)} \tag{4.3}$$

gdzie:

x^*	wartość parametru po modyfikacji,
$P^*(a_i\|\pi)$	prawdopodobieństwo warunkowe po normalizacji,
$P(a_i\|\pi)$	prawdopodobieństwo warunkowe przed normalizacją.

Wyznaczone współczynniki (c_0, c_1, d_0, d_1) opisują wpływ parametru x na prawdopodobieństwo a posteriori hipotezy $P(h|e)$ dla konkretnego kontekstu e. Dla ułatwienia analizy wyników i bezpośredniej oceny wrażliwości parametru x wprowadzono miarę wrażliwości [9]:

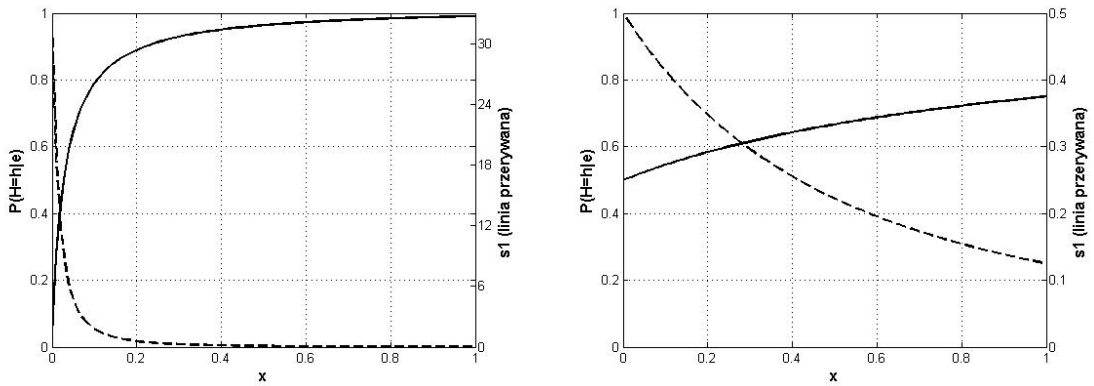
$$s1 = \left|\frac{d(P(h|e)(x))}{dx}\right| = \left|\frac{c_1(d_0 + d_1 x) - d_1(c_0 + c_1 x)}{(d_0 + d_1 x)^2}\right| \tag{4.4}$$

4.2.2. Interpretacja graficzna

Na rys.4.1 przedstawiono interpretację graficzną zależności (4.1) oraz współczynnika wrażliwości $s1$ (4.4) dla przykładowych danych. Na wykresie 4.1a widać, że parametr

przyjmujący wartość z zakresu od 0 do 0,1 jest bardzo wrażliwy, tzn. jego mała zmiana wartości powoduje duże zmiany prawdopodobieństwa hipotezy $P(H = h|e)$. Odzwierciedleniem tego jest wartość współczynnika wrażliwości $s1$ dla tego zakresu (od 3 do 32). Jeżeli natomiast wartość parametru jest z przedziału od 0,1 do 1 to parametr ten jest mało wrażliwy ($s1 < 3$). Na wykresie 4.1b przedstawiono sytuację, w której w całym zakresie zmian wartości parametru, wykazuje on małą wrażliwość ($s1 < 0,5$).

a) b)



Rys. 4.1. Interpretacja graficzna zależności (4.1) i współczynnika wrażliwości $s1$ (4.4): a) parametr wrażliwy w zakresie od 0 do 0,1, b) parametr mało wrażliwy w całym zakresie

Rozdział 5

Badania nad optymalizacją struktury sieci bayesowskiej

Optymalizacja polega na wyznaczeniu, spośród dopuszczalnych rozwiązań danego problemu, rozwiązania spełniającego przyjęte kryteria. W przypadku sieci bayesowskiej proces optymalizacji struktury można podzielić na dwa etapy:

etap 1 - modyfikowana jest struktura sieci, poprzez wprowadzenie dodatkowych gałęzi pomiędzy wybranymi węzłami a węzłami hipotez, w celu poprawy jakości wnioskowania.

etap 2 - zmniejszana jest liczba parametrów sieci, poprzez usunięcie gałęzi i/lub węzłów, które nie wpływają na jakość wnioskowania.

W etapie pierwszym dokonuje się wyboru węzłów, ze zbioru węzłów wyznaczonego na podstawie analizy wrażliwości węzłów obserwowalnych z wykorzystaniem warunkowej niezależności węzłów. Wyboru tego należy dokonać w oparciu o metody selekcji informacji [94, 95], tak aby zmienne reprezentowane przez wybrane węzły niosły jak najwięcej informacji o hipotezach. Wybrane węzły wykorzystywane są następnie w procesie tworzenia nowych gałęzi z węzłami hipotez. W etapie tym, liczba parametrów sieci bayesowskiej (prawdopodobieństw warunkowych) wzrasta. Należy zwrócić uwagę na to aby liczba parametrów nie zwiększyła się na tyle, że spowoduje to pogorszenie zdolności uogólniania modelowanego zagadnienia.

W drugim etapie dąży się do poprawy modelu, w sensie złożoności, jako czynnika wpływającego na przejrzystość reprezentacji graficznej modelu, oraz na szybkość wnioskowania i nakłady obliczeniowe w procesie wnioskowania. Wyznaczenie gałęzi i/lub węzłów do usunięcia odbywa się na podstawie ponownej analizy wrażliwości węzłów obserwowanych. Ze względu jednak na to, że sieci bayesowskie służą do reprezentacji wiedzy w warunkach niepewności, konieczna jest pewna ostrożność przy usuwaniu gałęzi i/lub węzłów. Usunięcie jednego z węzłów, może przy pewnych warunkach (niekompletne, niepewne, błędne dane), prowadzić do sytuacji, w której niemożliwe będzie prowadzenie wiarygodnego procesu wnioskowania. W przypadku, gdy brak jest jasnych przesłanek do usuwania gałęzi i/lub węzłów oraz gdy wzrost liczby parametrów nie powoduje pogorszenia zdolności uogólniania modelowanego zagadnienia wówczas etap drugi może zostać pominięty.

5.1. Identyfikowanie sieci bayesowskiej na podstawie danych symulacyjnych

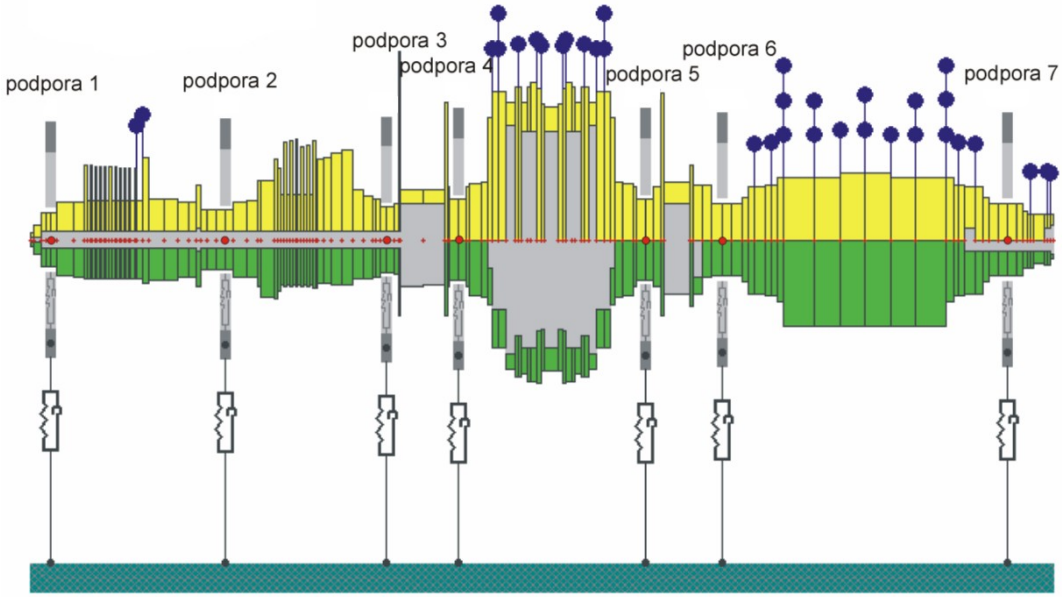
Weryfikacji zaproponowanej metody dokonano na sieci bayesowskiej utworzonej w wyniku procesu uczenia sieci z danych pochodzących z badań symulacyjnych.

5.1.1. Dane symulacyjne

W wyniku prac realizowanych w IMP PAN opracowano pakiet programów komputerowych tworzących spójne środowisko o nazwie MESWIR (**M**etoda **E**lementów **S**kończonych w analizie **WIR**ników), umożliwiające prowadzenie złożonych badań symulacyjnych wirników [58]. Środowisko MESWIR składa się z trzech zasadniczych modułów:

- zespołu programów opartych na modelach cieplnych łożysk ślizgowych,
- zespołu programów do analizy zagadnień kinostatyki układów wirnik-łożyska-fundament,
- zespołu programów do analizy zagadnień dynamiki układów wirnik-łożyska-fundament.

W ramach Projektu Badawczego Zamawianego K015/T10/2001 w IMP PAN, w oparciu o dokumentację techniczną, dane eksploatacyjne oraz dane pochodzące z systemu diagnostycznego DT200-1 [38], powstał model numeryczny turbogeneratora, pracującego w bloku 7 Elektrowni Kozienice, składającego się z turbiny 13K215 i generatora TWW-230-2L (rys. 5.1) [59]. Model ten uwzględnia wzajemnie sprzężone wymuszenia aerodynamiczne, mechaniczne i elektryczne oraz pozwala na generowanie relacji typu defekt→symptom.



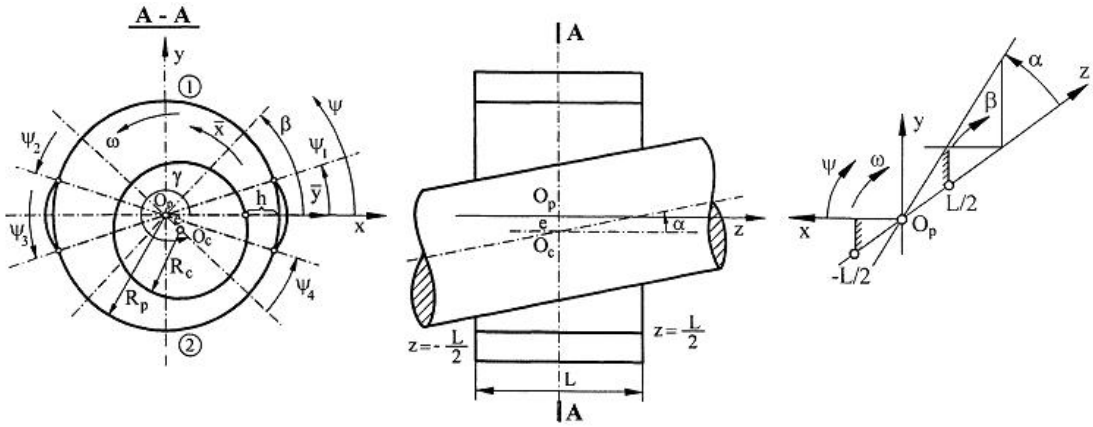
Rys. 5.1. Schemat modelu numerycznego turbogeneratora [59]

Udostępnione przez IMP PAN dane zawierają wyniki badań symulacyjnych przeprowadzonych dla:

- stanu bazowego - jeden stan [59],
- przekoszeń panwi - 28 stanów związanych z dwoma kątami przekoszenia, dwoma płaszczyznami oraz siedmioma podporami łożyskowymi, w których to przekoszenie wystąpiło [70],
- przemieszczeń pojedyńczych podpór - 56 stanów związanych z dwoma różnymi wartościami przemieszczeń w czterech kierunkach dla siedmiu podpór łożyskowych [91],
- przemieszczeń par podpór - 32 stany związane z jednoczesnym przemieszczaniem podpór 5 i 6 [91],
- pęknięć wału - 96 stanów związanych z pęknięciami w dwóch różnych punktach dla czterech przypadków obwodowego umiejscowienia pęknięcia oraz dla dwunastu wartości głębokość pęknięcia [6].

Spośród wymienionych symulowanych stanów do celów identyfikacji sieci przekonań wybrano trzy pierwsze klasy stanu (stan bazowy, przekoszenie panwi, przemieszczenie pojedynczej podpory), ponieważ reszta symulowanych klas stanowi dane niereprezentatywne, gdyż dotyczy tylko wybranych podpór łożyskowych (podpora 5 i 6) i miejsc pęknięć (w podporze 6 i pomiędzy podporą 6 i 7).

Identyfikowana sieć bayesowska miała za zadanie diagnozować przekoszenia pawi łożyskowych. Defekt typu przekoszenie panwi, polega na przekoszeniu osi panwi w stosunku do osi wału oraz, co za tym idzie, powoduje zmianę geometrii szczeliny smarnej łożyska. Model tego typu defektu został przedstawiony na rys. 5.2. Położenie i wielkość przekoszenia opisują dwa kąty α i β.



Rys. 5.2. Model przekoszenia panwi [58].

Tab. 5.1. Wartości kątów alfa dla płaszczyzny poziomej i pionowej

Nr	$\beta = 0°$		$\beta = 90°$	
łożyska	$\alpha(-)$	$\alpha(+)$	$\alpha(-)$	$\alpha(+)$
1	-0,1475	0,1475	-0,0675	0,0675
2	-0,0975	0,0975	-0,0625	0,0475
3	-0,095	0,095	-0,04	0,04
4	-0,0925	0,0925	-0,0625	0,0625
5	-0,1	0,1	-0,0425	0,0425
6	-0,0625	0,0625	-0,0225	0,0225
7	-0,0475	0,0475	-0,0425	0,0425

gdzie:

$\alpha(-)$ - wartość kąta α w kierunku ujemnym,

$\alpha(+)$ - wartość kąta α w kierunku dodatnim.

Oznaczenia:

x, y układ prostokątny związany ze środkiem panwi,

$\bar{x}$ współrzędna obwodowa,

$\bar{y}$ współrzędna po grubości filmu,

z współrzędna osiowa (po szerokości panwi),

α kąt określający przekoszenie czopa,

β kąt określający płaszczyznę w której następuje przekoszenie o kąt α,

e mimośrodowość (w linii środkowej łożyska),

γ kąt linii środków (w linii środkowej).

Defekt przekoszenia panwi został zasymulowany w płaszczyźnie poziomej (X-Z, $\beta = 0°$) i pionowej (Y-Z, $\beta = 90°$) dla każdego z siedmiu łożysk w kierunku dodatnim i ujemnym (wartość kąta α (tab.5.1)). Oznacza to, że badania symulacyjne zostały przeprowadzone dla czterech wartości defektu dla każdego łożyska (łącznie 28 przypadków).

5.1.2. Przygotowanie danych

W przypadku diagnostycznego systemu doradczego służącego do diagnozowania przekoszeń panwi łożyskowej, nie jest konieczne reprezentowanie przekoszenia za pomocą kątów α i β. Potencjalny użytkownik będzie raczej oczekiwał od systemu doradczego wnioskowania na temat tego czy występuje, czy też nie przekoszenie panwi i jeżeli tak to której. Dlatego też zdecydowano, że stan dla każdego przypadku reprezentowany będzie

jako wektor stanu

$$C = [c_1, \ldots, c_k, \ldots, c_7] \tag{5.1}$$

gdzie: c_k oznacza cechę stanu

$$c_k = \langle \text{"}\textit{przekoszenie k-tej panwi łożyskowej}\text{"}, \mathrm{TAK}, \mathrm{NIE} \rangle \tag{5.2}$$

Na podstawie badań literaturowych dotyczących diagnostyki wibroakustycznej oraz diagnozowania przekoszeń panwi [19,20,27,29,61,68,69,72,74,79,84,86,102] do budowy sieci przekonań wybrano następujące cechy sygnałów:

- wartość składowej widma amplitudowego o częstotliwości 1X
- wartość składowej widma amplitudowego o częstotliwości 2X
- faza składowej widma amplitudowego o częstotliwości 1X
- faza składowej widma amplitudowego o częstotliwości 2X

Cechy te wyznaczane są dla sygnałów przemieszczeń (drgań) czopa względem panwi, w płaszczyźnie promieniowej wału w kierunkach wzajemnie prostopadłych x i y (odpowiednio w kierunku poziomym i pionowym). Ze względu na cykliczność cechy jaką jest faza składowej 1X i 2X sygnału, jak również to, że w diagnozowaniu przekoszenia panwi, bardziej użyteczną cechą jest różnica faz sygnałów pochodzących z sąsiednich łożysk, postanowiono cechy te zastąpić współczynnikami korelacji liniowej Pearsona (5.3). Definicje cech zapisano przyjmując, że rozpatrywane sygnały zapisywane są w postaci szeregów czasowych $x(n)$ i $y(n)$ dla $n = 1 : N$ (odpowiednio dla kierunku poziomego i pionowego).

$$ccx = \frac{\sum_{n=1}^{N}(x^k(n) - \bar{x}^k)(x^{k+1}(n) - \bar{x}^{k+1})}{\sqrt{\sum_{n=1}^{N}(x^k(n) - \bar{x}^k)^2}\sqrt{\sum_{n=1}^{N}(x^{k+1}(n) - \bar{x}^{k+1})^2}} \qquad \text{dla} \quad k = 1:6 \tag{5.3}$$

$$ccy = \frac{\sum_{n=1}^{N}(y^k(n) - \bar{y}^k)(y^{k+1}(n) - \bar{y}^{k+1})}{\sqrt{\sum_{n=1}^{N}(y^k(n) - \bar{y}^k)^2}\sqrt{\sum_{n=1}^{N}(y^{k+1}(n) - \bar{y}^{k+1})^2}} \qquad \text{dla} \quad k = 1:6 \tag{5.4}$$

gdzie:

$x^k(n)$, $y^k(n)$	sygnały dla kierunku poziomego i pionowego pochodzące z k-tego łożyska,
$\bar{x}^k$, $\bar{y}^k$	wartość średnia sygnałów $x^k(n)$, $y^k(n)$.

W dalszej części rozprawy w zakresie identyfikowania sieci bayesowskiej na podstawie danych symulacyjnych przyjęto następujące oznaczenia cech sygnałów:

$lo(k).x.a1$	wartość składowej widma amplitudowego o częstotliwości 1X dla kierunku poziomego dla k-tego łożyska [m],
$lo(k).y.a1$	wartość składowej widma amplitudowego o częstotliwości 1X dla kierunku pionowego dla k-tego łożyska [m],
$lo(k).x.a2$	wartość składowej widma amplitudowego o częstotliwości 2X dla kierunku poziomego dla k-tego łożyska [m],
$lo(k).y.a2$	wartość składowej widma amplitudowego o częstotliwości 2X dla kierunku pionowego dla k-tego łożyska [m],
$lo(k,k+1).ccx$	wartość współczynnika korelacji liniowej pomiędzy sygnałami dla kierunku poziomego dla sąsiednich łożysk k i $k+1$,
$lo(k,k+1).ccy$	wartość współczynnika korelacji liniowej pomiędzy sygnałami dla kierunku pionowego dla sąsiednich łożysk k i $k+1$.

W wyniku tak przeprowadzonych działań uzyskano dane D składające się z 28 przypadków przekoszenia panwi łożyskowej, 56 przypadków przemieszczeń pojedynczej podpory oraz 1 przypadku bazowego (85 przypadków). Przypadki te reprezentowane są przez 40 cech sygnałów i 7 cech stanu. Dla przypadków przemieszczeń pojedynczej podpory i przypadku bazowego wszystkie cechy stanu przyjmują wartość NIE, gdyż w badaniach symulacyjnych dla tych przypadków nie był zakładany defekt przekoszenia panwi łożyskowej.

Generowanie dodatkowych przykładów

Dane pochodzące z badań symulacyjnych uzyskiwane są w wyniku obliczeń przeprowadzanych za pomocą oprogramowania komputerowego z zastosowaniem odpowiednich modeli numerycznych [28, 58]. W większości przypadków badania takie pozwalają uzyskiwać dane praktycznie dokładne w tym sensie, że odchyłki modelu są dużo mniejsze w stosunku do odchyłek, które wynikałyby z rzeczywistych pomiarów. Dlatego też, w celu uzyskania danych zbliżonych do danych pomiarowych oraz zwiększenia liczby przypadków uczących, zdecydowano się na wygenerowanie dodatkowych przykładów (dziewięć dla każdego przypadku pochodzącego z badań symulacyjnych) ze śztucznie" wprowadzonymi zakłóceniami ($85 + 85 \cdot 9 = 850$ przykładów).

W przypadku rozważanych danych wielkość wprowadzonych zakłóceń bazuje na analizie wpływu szumów pomiarowych na wartości sygnałów przemieszczeń (wynikających np. z niejednorodności geometrycznej, magnetycznej, elektrycznej, wpływu temperatury, ciśnienia itp. [29]). Na tej podstawie zakłócenia zostały wygenerowane dla każdej cechy sygnału jako liczby pseudolosowe o rozkładzie normalnym ze średnią 0 i wariancją równą 1% wartości cechy sygnału przypadku z badań symulacyjnych.

Dyskretyzacja cech sygnałów

Ze względu na dostępność informacji o stanie, proces dyskretyzacji ciągłych wartości cech sygnałów został przeprowadzony na podstawie metody minimalnej długości kodu (MDLP - Minimal Description Length Principle). Jest to metoda nadzorowana oparta

na mierze entropii z warunkiem stopu bazującym na zasadzie minimalnej długości kodu. Metoda MDLP została przedstawiona w rozdz. 3.2.1 i dotyczy dyskretyzacji w oparciu o jedną cechę stanu. W przypadku posiadanych danych mamy do czynienia z wektorem stanu (7 cechami stanu), dlatego też metoda dyskretyzacji została zmodyfikowana. Modyfikacja polegała na tym, że punktem binarnej dyskretyzacji przestrzeni wartości cechy sygnału jest próg, dla którego wartość ważonej entropii, liczona dla wszystkich 7 cech stanów osobno, jest najmniejsza.

Wybór tej metody został podyktowany tym, że metoda ta jest jedną z najlepszych i najwydajniejszych metod dyskretyzacji [37, 66, 89], jak również jest ona skutecznie stosowana w procesie uczenia sieci przekonań [43].

Wyniki procesu dyskretyzacji wartości cech sygnałów przedstawiono w [8].

5.1.3. Uczenie sieci bayesowskiej

Uczenie struktury sieci bayesowskiej, w oparciu o dane pochodzące z badań symulacyjnych, wykonano z zastosowaniem algorytmu GES (Greedy Equivalence Search - rozdz. 3.2.3), zakładając jako miarę oceny sieci miarę BIC. W przypadku uczenia parametrów wykorzystano podejście bayesowskie zakładając hiperparametry $\alpha_{ijk} = \alpha/(r_i \cdot q_i)$ [15], gdzie równoważny rozmiar przykładów α przyjęto jako 1 (rozdz. 3.2.2). Cały proces uczenia sieci bayesowskiej przeprowadzono w środowisku MATLAB [99], z wykorzystaniem pakietów narzędziowych BNT (Bayes Net Toolbox for Matlab [7]) i BNT-SLP (BNT Structure Learning Package [11]).

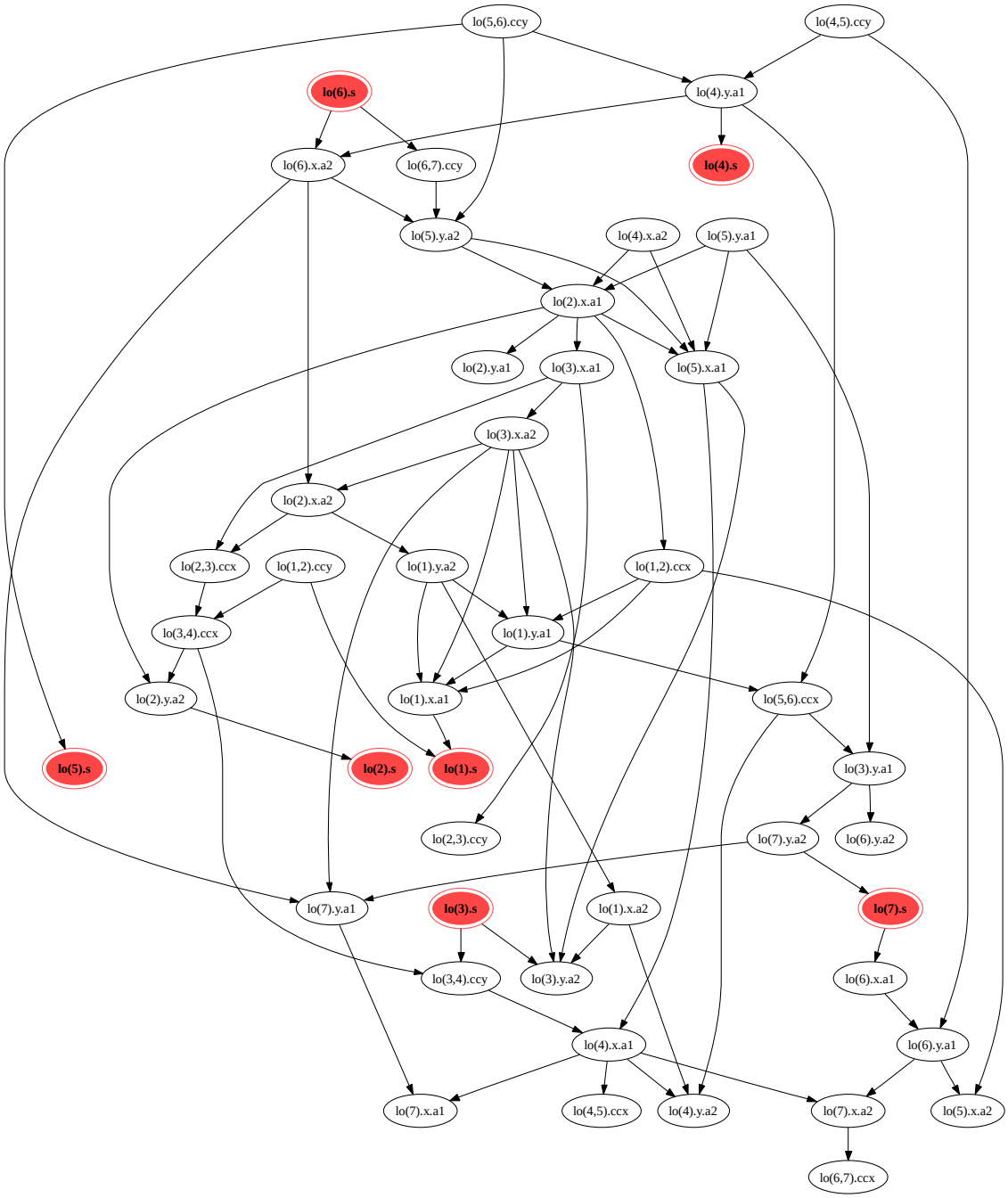
Wyboru algorytmów uczenia sieci bayesowskiej, dokonano w możliwie jak najszerszym aspekcie, w oparciu o analizę wad i zalet [3, 65] biorąc pod uwagę ich przydatność w procesie identyfikacji modeli diagnostycznych.

W wyniku uczenia sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych uzyskano sieć, której struktura została przedstawiona na rys.5.3, gdzie kolorem ■ zaznaczono węzły hipotez, oznaczone jako $lo(k).s$ - przekoszenie k-tej panwi łożyskowej.

5.1.4. Weryfikacja sieci bayesowskiej

W praktyce zwykle dąży się do tego, aby modele diagnostyczne reprezentowały dziedzinę problemu w jak największym stopniu uogólniając analizowany problemem diagnostyczny. Nie chodzi więc o model, w wyniku działania którego otrzymujemy maksymalnie zgodne wyniki z danymi na podstawie których został on wyznaczony, ale o to by znaleźć model, który będzie poprawnie wnioskował dla każdych innych danych z dziedziny zastosowania modelu. Chodzi zatem o to, by model diagnostyczny posiadał zdolność uogólniania problemu.

Estymację błędu sieci bayesowskiej otrzymanej w wyniku uczenia dokonano na podstawie metody *rotacji (leave-k-out)*. Metoda ta uznawana jest jako jedna z lepszych metod estymacji ze względu na dużą liczebność zbioru uczącego oraz wiarygodne wyniki [47, 58]. Polega ona na podziale N przykładów na dwa rozłączne zbiory, przy czym



Rys. 5.3. Struktura sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych

zbiór testujący zawiera k przykładów podczas gdy uczący pozostałe $N - k$ przykładów. Uczenie i testowanie prowadzi się m-krotnie (gdzie: $m = N/k$). W przypadku rozpatrywanej sieci bayesowskiej, zbiór testujący przyjmowano jako jeden przypadek pochodzącego z badań symulacyjnych wraz z dziewięcioma dodatkowo wygenerowanymi przykładami ($k = 10$). Ponadto, dla każdego zbioru danych uczących przeprowadzano osobno dyskretyzację cech ciągłych. Przedział pomiędzy wartościami błędu estymacji, wyznaczonymi metodą leave-k-out i metodą resubstytucji (testowanie odbywa się za pomocą danych uczących), pozwala określić możliwości uogólniania modelu.

W przypadku sieci bayesowskich do oceny jakości wnioskowania należy również przyjąć próg (progi) p, który możemy interpretować jako maksymalne ryzyko błędu, jakie jesteśmy skłonni zaakceptować. Wybór wartości p zależy od natury problemu i od tego jak dokładnie będą weryfikowane wyniki. W przypadku sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych przyjęto progi $p_1 = 0,4$ i $p_1 = 0,6$ tzn:

$P(lo(k).s = val(c_k)) > 0,6$ hipoteza *przekoszenie k-tej panwi łożyskowej*=$val(x_i)$ uznawana jest za prawdziwą,

$0,4 \leq P(lo(k).s = val(c_k)) \leq 0,6$ nie można potwierdzić ani odrzucić hipotezy *przekoszenie k-tej panwi łożyskowej*=$val(c_k)$

$P(lo(k).s = val(c_k)) < 0,4$ hipoteza *przekoszenie k-tej panwi łożyskowej*=$val(c_k)$ uznawana jest za nieprawdziwą

gdzie: $val(c_k)$ jest wartością {TAK, NIE} cechy stanu *"przekoszenie k-tej panwi łożyskowej"* i wiadomo, że $lo(k).s = val(c_k)$.

W rozpatrywanej sieci mamy do czynienia z przykładami, w których stan jest reprezentowany za pomocą wektora cech stanu opisujących przekoszenia kolejnych panwi łożyskowych. Dlatego też do oceny jakości wnioskowania przyjęto, że przykład jest:

- błędnie rozpoznawany, jeżeli co najmniej jedna hipoteza uznawana jest za nieprawdziwą,
- niesklasyfikowany, jeżeli brak jest hipotez uznawanych za błędne i co najmniej jednej hipotezy nie można potwierdzić ani odrzucić,
- poprawnie rozpoznawany, jeżeli wszystkie hipotezy uznawane są za prawdziwe.

Wyniki weryfikacji metodą rotacji (leave-k-out) i metodą resubstytucji sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych, przedstawiono w formie tabel (tab. 5.2 i 5.3) zawierających wartości procentowe liczby hipotez uznanych za nieprawdziwe i hipotez, których nie można potwierdzić ani odrzucić dla kolejnych panwi łożyskowych. W tabelach pokazano również procent *przykładów* błędnie sklasyfikowanych (odniesiony do liczby wszystkich przykładów - 850 przykładów). Uwzględniono przy tym:

- błąd I rodzaju - przykład uznawano za błędnie sklasyfikowany, jeżeli odrzucono co najmniej jedną z siedmiu hipotez *"przekoszenie k-tej panwi łożyskowej"*=*TAK*, która w rzeczywistości jest prawdziwa,
- błąd II rodzaju - przykład uznawano za błędnie sklasyfikowany, jeżeli uznano co najmniej jedną z siedmiu hipotez *"przekoszenie k-tej panwi łożyskowej"*=*TAK* jako prawdziwą, która w rzeczywistości jest fałszywa oraz nie wystąpił błąd I rodzaju.
- średni błąd w którym przyjęto równe koszty dla błędu I i II rodzaju.

Wyniki weryfikacji dla każdego przykładu przedstawiono w [8].

Tab. 5.2. Wyniki weryfikacji metodą resubstytucji sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych

	% przykładów niepoprawnie rozpoznanych	% przykładów niesklasyfikowanych
lo(1).s=TAK	5	0
lo(1).s=NIE	0	0
lo(1).s={TAK, NIE}	0,2	0
lo(2).s=TAK	100	0
lo(2).s=NIE	0	0
lo(2).s={TAK, NIE}	4,7	0
lo(3).s=TAK	2,5	2,5
lo(3).s=NIE	0,6	1,4
lo(3).s={TAK, NIE}	0,7	1,4
lo(4).s=TAK	50	0
lo(4).s=NIE	0,0	0
lo(4).s={TAK, NIE}	2,4	0
lo(5).s=TAK	100	0
lo(5).s=NIE	0	0
lo(5).s={TAK, NIE}	4,7	0
lo(6).s=TAK	5	0
lo(6).s=NIE	0,0	1,2
lo(6).s={TAK, NIE}	0,2	1,2
lo(7).s=TAK	0	0
lo(7).s=NIE	0	0
lo(7).s={TAK, NIE}	0	0
Błąd I rodzaju	37,5	0,4
Błąd II rodzaju	1	1,2
Błąd średni	19	0,8

Na podstawie otrzymanych wyników weryfikacji stwierdzono, że jakość sieci bayesowskiej, do detekcji uszkodzenia jakim jest przekoszenie panwi łożyskowych, jest zadowalająca. Średni błąd, dla weryfikacji metodą resubstytucji, przykładów niepoprawnie rozpoznanych wynosi 19%, a dla przykładów niesklasyfikowanych $0,8\%$. W przypadku weryfikacji metodą leave-k-out, błędy te wynoszą odpowiednio 18% i $6,3\%$. Ponadto, porównując wyniki weryfikacji metodą resubstytucji i rotacji [8] można zauważyć, że nie następuje znaczne pogorszenie jakości rozpatrywanej sieci (pogorszenie jakości widać

Tab. 5.3. Wyniki weryfikacji metodą rotacji (leave-k-out) sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych

	% przykładów niepoprawnie rozpoznanych	% przykładów niesklasyfikowanych
lo(1).s=TAK	10	0
lo(1).s=NIE	1,2	0
lo(1).s={TAK, NIE}	1,6	0
lo(2).s=TAK	52,5	47,5
lo(2).s=NIE	1,2	0
lo(2).s={TAK, NIE}	3,6	2,2
lo(3).s=TAK	7,5	0
lo(3).s=NIE	0,2	1,1
lo(3).s={TAK, NIE}	0,6	1,1
lo(4).s=TAK	50	0
lo(4).s=NIE	0,1	0
lo(4).s={TAK, NIE}	2,5	0
lo(5).s=TAK	100	0
lo(5).s=NIE	0	4,9
lo(5).s={TAK, NIE}	4,7	4,7
lo(6).s=TAK	5	0
lo(6).s=NIE	2,5	0
lo(6).s={TAK, NIE}	2,6	0
lo(7).s=TAK	0	0
lo(7).s=NIE	0	1,2
lo(7).s={TAK, NIE}	0	1,2
Błąd I rodzaju	32,1	6,8
Błąd II rodzaju	4	5,8
Błąd średni	18	6,3

po wzroście liczby przykładów niesklasyfikowanych). W związku z tym, można wnioskować, że jakość sieci bayesowskiej, nauczonej na wszystkich przykładach, jest najlepsza i dodatkowo sieć ta posiada dużą zdolność do uogólniania zdobytej w trakcie uczenia wiedzy. W dalszych działaniach dotyczących optymalizacji struktury sieci bayesowskich, wykorzystywana więc będzie sieć nauczona na wszystkich przykładach.

5.2. Wyniki analizy wrażliwości węzłów obserwowanych

Analizę wrażliwości węzłów obserwowanych sieci bayesowskiej do diagnozowania przekoszenia panwi łożyskowych przeprowadzono w oparciu o analizę warunkowej niezależności pomiędzy węzłami obserwowanymi i węzłami hipotez. Należy zwrócić uwagę, że wyniki takiej analizy zależą od kontekstu (ustalonych wartości węzłów obserwowanych). Wynika to z własności warunkowej niezależności węzłów. W przypadku rozpatrywanej sieci dla wszystkich przypadków kontekst jest stały. W [8] przedstawiono listę węzłów, dla których zmiany wartości nie mają wpływu na odpowiednie węzły hipotez (węzły warunkowo niezależne w stosunku do węzłów hipotez).

5.3. Optymalizacja struktury

Modyfikacji struktury rozpatrywanej sieci bayesowskiej, której zadaniem było polepszenie jakości wnioskowania, dokonano poprzez wprowadzenie dodatkowych gałęzi pomiędzy wybranymi węzłami obserwowanymi a węzłami hipotez. Wyboru węzłów dokonano w oparciu o wcześniej przeprowadzoną analizę wrażliwości oraz metodę selekcji wykorzystującą teorię informacji [94, 95]. Analiza wrażliwości węzłów obserwowanych pozwoliła wyznaczyć zbiory węzłów obserwowanych, których zmiana wartości nie powodowała zmian prawdopodobieństwa a posteriori odpowiednich węzłów hipotez. Wykorzystując miarę ilości informacji (5.5) z wyznaczonych zbiorów węzłów obserwowanych wybrano po dwa węzły dla każdego węzła hipotez, które niosą łącznie najwięcej informacji o hipotezach w tych węzłach.

$$I(H:X) = Ent(H) - Ent(H|X) \tag{5.5}$$

gdzie:

$I(H:X)$ ilość informacji dostarczona przez węzeł X o hipotezach w węźle H,

$Ent(H)$ entropia hipotez w węźle H wyrażona jako:

$$Ent(H) = -\sum_{i=1}^{n} P(H=h_i) log_2 P(H=h_i) \tag{5.6}$$

$Ent(H|X)$ średnia entropia warunkowa hipotez w węźle H w zbiorowościach wyróżnionych ze względu na węzeł X, definiowana jako:

$$Ent(H|X) = -\sum_{x \in X} \sum_{i=1}^{n} P(H=h_i, X=x) log_2 P(H=h_i|X=x) \tag{5.7}$$

Podstawiając (5.6) i (5.7) do (5.5) można ilość dostarczonej informacji wyrazić wzorem:

$$I(H:X) = \sum_{i=1}^{n} \sum_{x \in X} P(X=x|H=h_i) P(H=h_i) log_2 \frac{P(X=x|H=h_i)P(H=h_i)}{P(H=h_i)P(X=x)} \tag{5.8}$$

W tabeli 5.4 przedstawiono węzły obserwowane wybrane ze zbiorów węzłów przedstawionych w [8], które zostały wyznaczone na podstawie metody selekcji z wykorzystaniem miary ilości informacji $I(H : X)$ i łącznie niosą najwięcej informacji o hipotezach odpowiednich węzłów hipotez.

Tab. 5.4. Węzły wybrane do utworzenia dodatkowych gałęzi w sieci

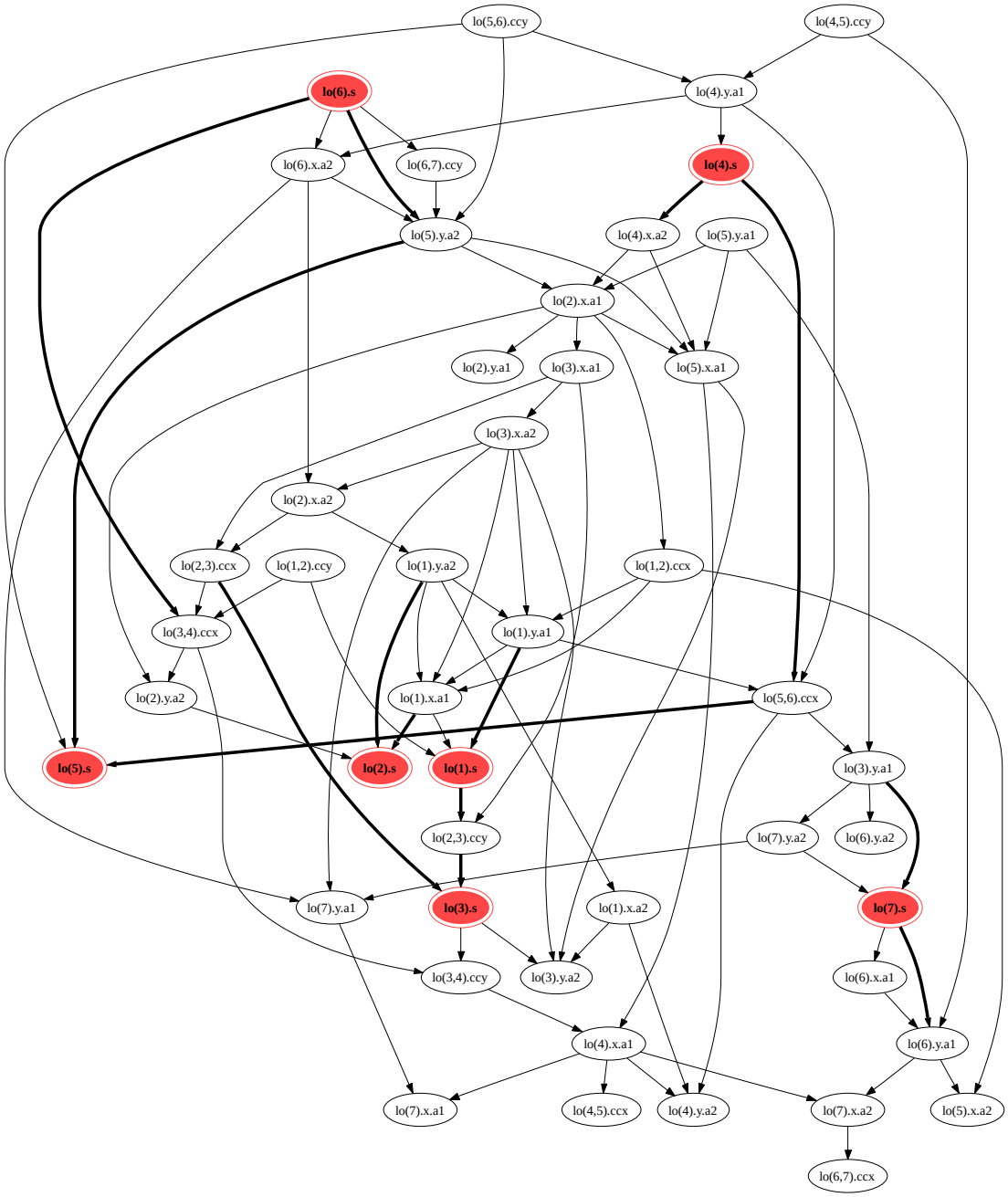
Węzły hipotez	Wybrane węzły
lo(1).s	lo(1).y.a1 lo(2,3).ccy
lo(2).s	lo(1).x.a1 lo(1).y.a2
lo(3).s	lo(2,3).ccx lo(2,3).ccy
lo(4).s	lo(4).x.a2 lo(5,6).ccx
lo(5).s	lo(5).y.a2 lo(5,6).ccx
lo(6).s	lo(3,4).ccx lo(5).y.a2
lo(7).s	lo(3).y.a1 lo(6).y.a1

Na podstawie otrzymanych wyników utworzono dodatkowe gałęzie w rozpatrywanej sieci bayesowskiej. Kierunek gałęzi wyznaczono tak, aby ilość parametrów sieci była jak najmniejsza. Na etapie tym dodatkowo należało wziąć pod uwagę warunek, wynikający z definicji sieci bayesowskiej (definicja 2.1, str. 15), że sieć bayesowska musi być grafem acyklicznym. Nowe gałęzie przedstawiono na rys.5.4, w postaci pogrubionych linii. Kolorem ■ zaznaczono węzły hipotez (oznaczone jako $lo(k).s$ - przekoszenie k-tej panwi łożyskowej).

Brakujące parametry, o które zwiększyła się rozpatrywana sieć bayesowska w wyniku dodania nowych gałęzi, wyznaczono w procesie ponownego uczenia ilościowego. Dodatkowo, w celu sprawdzenia czy model w postaci sieci bayesowskiej, zbytnio nie dopasował się do danych uczących, przeprowadzono weryfikację analizowanej sieci metodami resubstytucji i leave-k-out (opisanymi w rozdz.5.1.4). Wyniki weryfikacji przedstawiono w tabelach 5.5 i 5.6 oraz w [8].

Jakość wnioskowania rozpatrywanej sieci bayesowskiej znacznie wzrosła (tab.5.2 i 5.5), o czym świadczy zmniejszenie, w dużym stopniu, błędów klasyfikacji. W tym przypadku błąd średni klasyfikacji zmniejszył się z 19% na 6%, w wyniku zmniejszenia się błędu I rodzaju (z 37,5% na 8,9%). Błąd II rodzaju wzrósł z 1% na 3%, niemniej jednak pozostał na niskim poziomie. Procent przykładów niesklasyfikowanych natomiast nie zmienił się znacznie, lecz już przed procesem optymalizacją był stosunkowo niski (0,8%).

Na podstawie otrzymanych wyników stwierdzono, że nie ma potrzeby przeprowadza-



Rys. 5.4. Struktura sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych z dodatkowymi gałęziami między węzłami

nia drugiego etapu optymalizacji, jakim jest usunięcie węzłów, wyznaczonych w ponownej analizie wrażliwości węzłów obserwowanych (węzłów których zmiana wartości nie powoduje zmian prawdopodobieństw a posteriori w węzłach hipotez). Wynika to z faktu, że oszacowana jakość modelu jest w pełni zadowalająca, a otrzymana, w wyniku optymalizacji, sieć bayesowska nie wykazuje cech nadmiernego dopasowania się do danych (tab.5.5 i 5.6. Dodatkowo, nie istnieją żadne przesłanki związane z przejrzystością reprezentacji graficznej modelu, szybkością wnioskowania oraz nakładami obliczeniowymi w procesie wnioskowania, które potwierdzałyby potrzebę wykonania tego etapu.

Tab. 5.5. Wyniki weryfikacji metodą resubstytucji sieci bayesowskiej (z dodatkowymi gałęziami) do diagnozowania przekoszeń panwi łożyskowych

	% przykładów niepoprawnie rozpoznanych	% przykładów niesklasyfikowanych
lo(1).s=TAK	5	0
lo(1).s=NIE	0	0
lo(1).s={TAK, NIE}	0,2	0
lo(2).s=TAK	0	0
lo(2).s=NIE	0	1,2
lo(2).s={TAK, NIE}	0,0	1
lo(3).s=TAK	2,5	0
lo(3).s=NIE	0,0	0,0
lo(3).s={TAK, NIE}	0,1	0,0
lo(4).s=TAK	50	0
lo(4).s=NIE	0,0	0
lo(4).s={TAK, NIE}	2,4	0
lo(5).s=TAK	0	0
lo(5).s=NIE	2	0
lo(5).s={TAK, NIE}	1,5	0
lo(6).s=TAK	5	0
lo(6).s=NIE	0,0	1,6
lo(6).s={TAK, NIE}	0,2	1,5
lo(7).s=TAK	0	0
lo(7).s=NIE	0	0
lo(7).s={TAK, NIE}	0	0
Błąd I rodzaju	8,9	0,0
Błąd II rodzaju	3	1,2
Błąd średni	6	0,6

5.4. Podsumowanie i wnioski

W rozdziale przedstawiono proces optymalizacji struktury sieci bayesowskiej do diagnozowania przekoszeń panwi łożyskowych turbogeneratora, w oparciu o analizę wrażliwości węzłów obserwowanych. Rozpatrywaną sieć bayesowską, wykorzystaną w celu zobrazowania takiego podejścia, otrzymano w wyniku procesu uczenia sieci na podstawie danych pochodzących z badań symulacyjnych. Dodatkowo, w celu identyfikacji modelu w postaci sieci bayesowskiej, należało przeprowadzić dyskretyzację cech ciągłych, co zostało wykonane za pomocą metody minimalnej długości kodu. Zwiększono również liczbę przykładów uczących wprowadzając dodatkowe zakłócenia, mające na celu uzyskanie danych o charakterze zbliżonym do danych pomiarowych.

Na podstawie uzyskanych wyników wykazano skuteczność zaproponowanej metody optymalizacji struktury sieci bayesowskiej. Otrzymana w wyniku procesu optymaliza-

Tab. 5.6. Wyniki weryfikacji metodą rotacji (leave-k-out) sieci bayesowskiej (z dodatkowymi gałęziami) do diagnozowania przekoszeń panwi łożyskowych

	% przykładów niepoprawnie rozpoznanych	% przykładów niesklasyfikowanych
lo(1).s=TAK	10	0
lo(1).s=NIE	0	0
lo(1).s={TAK, NIE}	0,5	0
lo(2).s=TAK	75	0
lo(2).s=NIE	0	0,2
lo(2).s={TAK, NIE}	3,5	0
lo(3).s=TAK	22,5	0
lo(3).s=NIE	1,5	0,0
lo(3).s={TAK, NIE}	2,5	0,0
lo(4).s=TAK	50	0
lo(4).s=NIE	0,1	0
lo(4).s={TAK, NIE}	2,5	0
lo(5).s=TAK	30	0
lo(5).s=NIE	2	1
lo(5).s={TAK, NIE}	2,9	1
lo(6).s=TAK	25	2,5
lo(6).s=NIE	3,5	0,0
lo(6).s={TAK, NIE}	4,5	0,1
lo(7).s=TAK	15	0
lo(7).s=NIE	0	1
lo(7).s={TAK, NIE}	1	1
Błąd I rodzaju	32,5	0,4
Błąd II rodzaju	7	1,3
Błąd średni	20	0,8

cji sieć bayesowska do diagnozowania przekoszenia panwi łożyskowych, wykazywała się znacznie lepszą jakością wnioskowania, niż sieć, która powstała wyłącznie w wyniku procesu uczenia sieci na podstawie danych.

Szczegółowa analiza wyników weryfikacji otrzymanych sieci bayesowskich [8] pokazuje, że zagadnienie diagnozowania przekoszeń panwi łożyskowych jest złożonym zadaniem. Wynika to z faktu, że przemieszczenia podpór łożyskowych, mogą wywoływać skutki zbliżone do tych, które powstają przy przekoszeniu sąsiadujących z podporą, panwi łożyskowych. Dlatego też, należy pamiętać, że rozpatrywanie wyłącznie cech stanu dotyczących przekoszenia panwi łożyskowych, może prowadzić do mylnych wniosków dotyczących stanu badanego obiektu. Rozwiązaniem tego problemu, może być rozszerzenie zbioru rozpatrywanych cech stanu o cechy opisujące położenie podpór. Należy jednak mieć na uwadze, że to rozszerzenie może nie przynieść oczekiwanych rezultatów, z uwagi na pojawiający się problem rozróżnialności niesprawności.

Rozdział 6

Badania nad optymalizacją parametrów sieci bayesowskiej

Określenie parametrów sieci przekonań (prawdopodobieństw warunkowych) w większości przypadków bazuje na wiedzy i intuicji ekspertów z danej dziedziny. Często jednak wiedza w zakresie eksploatacji maszyn i urządzeń jest niejednoznaczna, niepełna, a nawet sprzeczna [77,78]. W takim przypadku założono, że wśród dwóch sieci bayesowskich A i B, lepszą siecią jest ta, która ma mniejszą maksymalną wrażliwość parametrów. Z tego założenia wynika, że błąd (niepewność) parametrów sieci bayesowskiej, dla której maksymalna wrażliwość jest mniejsza, skutkuje mniejszym błędem w węzłach hipotez. Bazując na tym podejściu, można rozważać następujące zadania optymalizacji sieci bayesowskiej:

- poszukiwanie węzłów z największą wrażliwością (największym wpływem na węzły hipotez) i uważna kontrola (dostrajanie) parametrów tych węzłów,
- modyfikacja sieci w celu uzyskania zbliżonych do siebie wrażliwości węzłów, gwarantujących ich jednakowy udział w procesie wnioskowania (wpływ na węzły hipotez).

6.1. Badana sieć bayesowska

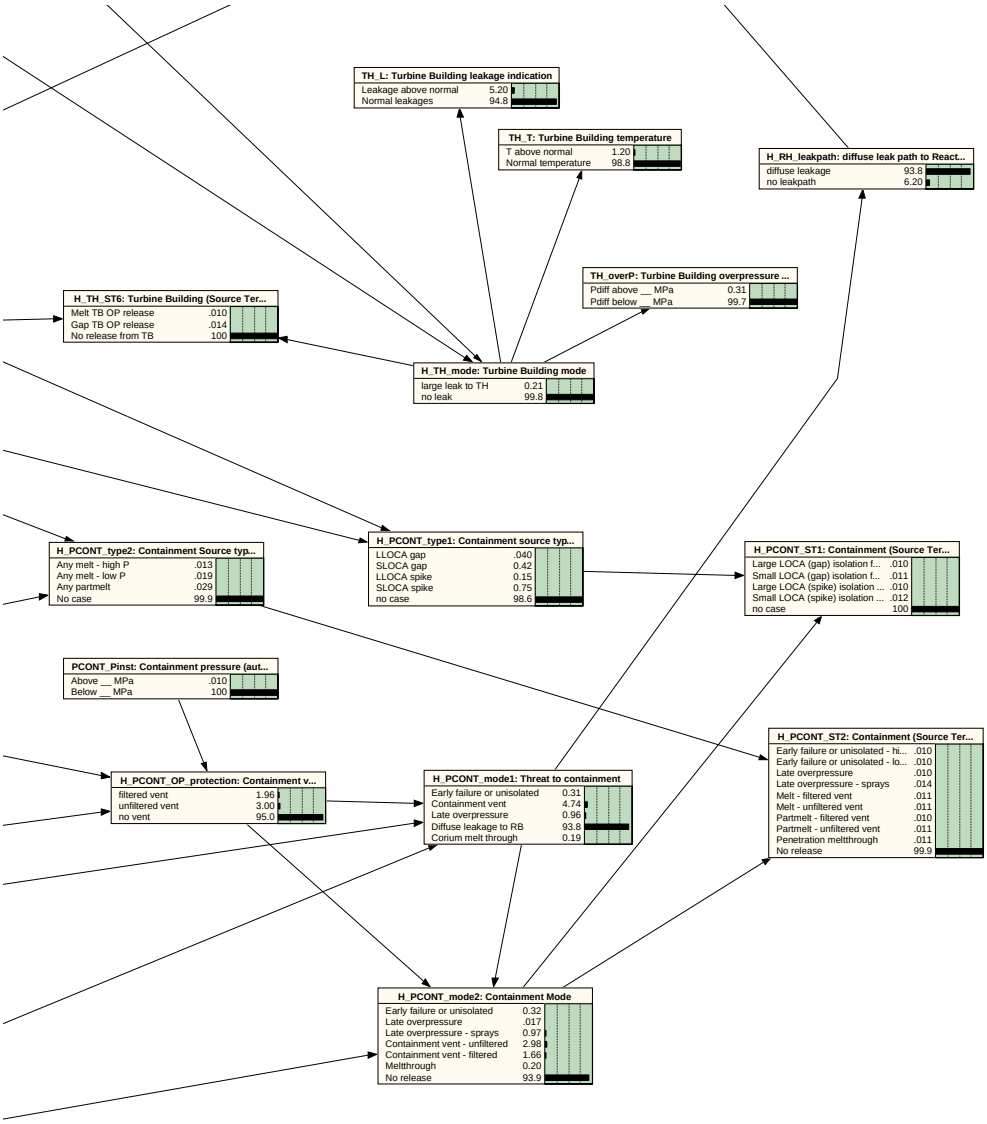
Badana sieć bayesowska została utworzona w ramach projektu STERPS (Source Term Indicator Based on Plant Status) w 5 Projekcie Ramowym Unii Europejskiej Euroatom (European Union 5-th Euroatom Framework) [46]. Celem tego projektu było stworzenie narzędzia, na bazie sieci bayesowskiej, umożliwiającego na podstawie kluczowych obserwacji szybkie diagnozowanie stanu elektrowni jądrowej, a następnie ocenę ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii. W ramach projektu stworzono ogólną sieć bayesowską (generic bayesian network). Sieć ta stanowiła punkt wyjściowy w procesie tworzenia konkretnych sieci

---

Badania nad optymalizacją parametrów sieci bayesowskiej były wykonywane w ramach projektu STERPS (Source Term Indicator Based on Plant Status) w 5 Projekcie Ramowym Unii Europejskiej Euroatom (European Union 5-th Euroatom Framework) w zespole kierowanym przez Prof. dr hab. inż. Wojciecha Cholewę [9].

dopasowanych do różnych konstrukcji reaktorów jądrowych (reprezentujących główne konstrukcje działających w Europie elektrowni jądrowych).

Rozpatrywana sieć bayesowska dotyczy szwedzkiego reaktora wrzącego Oskarshamn 3 (rys.6.1) [41] . Proces dopasowywania tej sieci bazował na wiedzy ekspertów oraz na opisach działających systemów jak również na analizach odpowiedzi tych systemów. Dodatkowo, w celu ustalenia parametrów sieci, rozpatrywane były probabilistyczne analizy bezpieczeństwa (PSA - Probabilistic Safety Analysis), analizy cieplno-przepływowe [56] oraz analizy groźnych awarii (severe accident analysis) [39]. W wielu przypadkach parametry sieci przyjmowane były w oparciu o subiektywne oszacowania ekspertów. Dotyczyło to głównie węzłów reprezentujących złożone zjawiska (np. prawdopodobieństwa wybuchu wodoru) i węzłów z dużą liczbą kombinacji wartości węzłów nadrzędnych (rodziców).



Rys. 6.1. Fragment sieci bayesowskiej dla reaktora Oskarshamn 3 [41]

6.2. Analiza wrażliwości parametrów

Analiza wrażliwości parametrów sieci bayesowskiej dla reaktora Oskarshamn 3 została przeprowadzona zgodnie z metodologią przedstawioną w rozdziale 4.2. Ze względu na fakt, że rozpatrywana sieć została stworzona z wykorzystaniem programu NETICA [85], dlatego też dla celów prowadzenia analizy wrażliwości parametrów opracowany został program w środowisku Visual C++ [73], z wykorzystaniem interfejsu programowania aplikacji NETICA API [83].

6.2.1. Plan badań

W przypadku prowadzenia analizy wrażliwości parametrów sieci bayesowskiej najlepszą sytuacją byłaby taka, w której znane są:

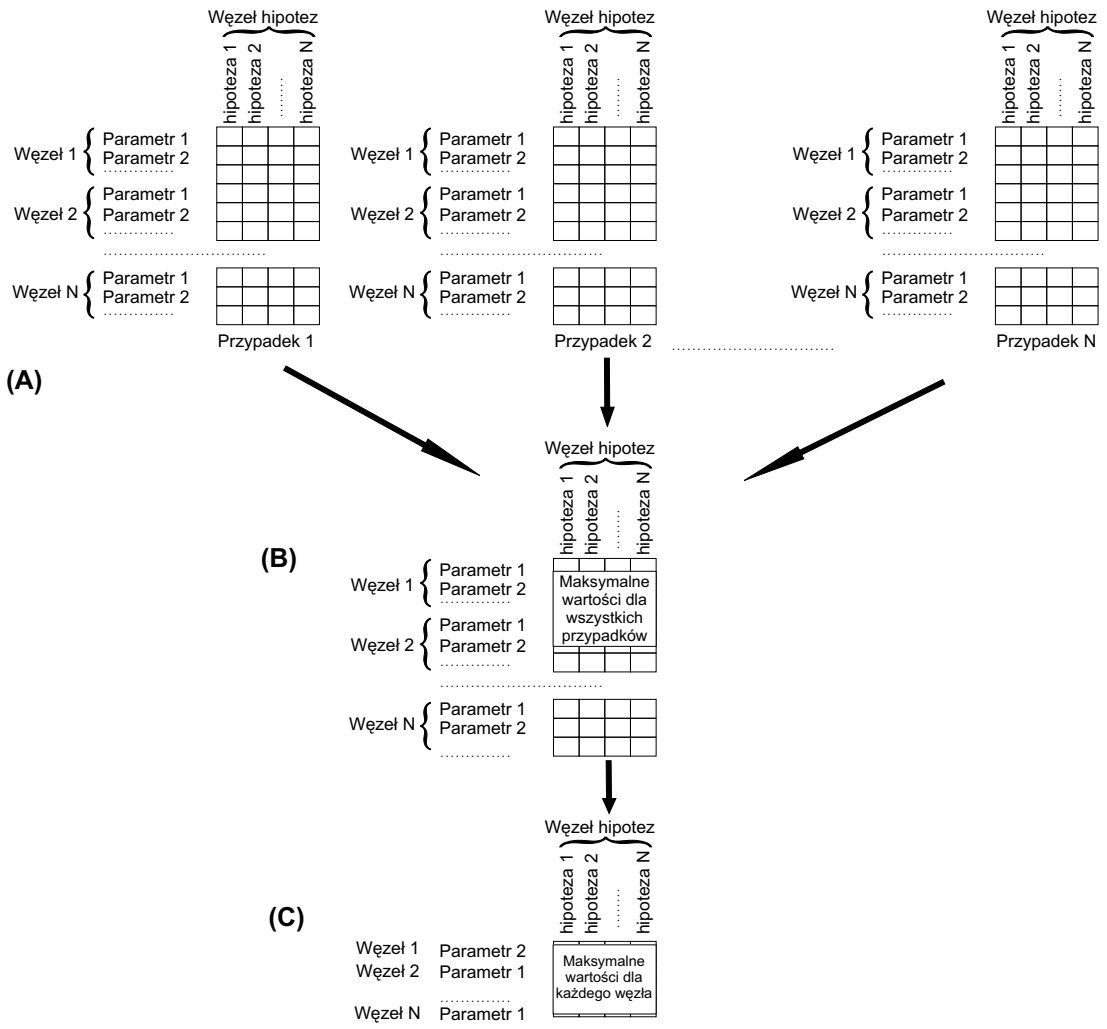
- struktura sieci,
- parametry sieci,
- odpowiednia liczba przypadków (kontekstów).

W rozpatrywanym przypadku, znane są tylko struktura i parametry sieci bayesowskiej. W związku z tym, przypadki (konteksty) zostały wygenerowane losowo, jako kombinacja wartości węzłów obserwowanych (dostępna była informacja, które węzły są węzłami obserwowanymi). Ponadto przyjęto, że każdy węzeł obserwowany posiada dodatkową wartość *ńieznana"* (wartość węzła nie jest ustalana), zgodnie z założeniem, że sieć bayesowska reprezentuje wiedzę w warunkach niepewności. W wyniku tego procesu otrzymano 200 przypadków (kontekstów) dla celów wyznaczenia wrażliwości parametrów rozpatrywanej sieci. Jako węzeł hipotez przyjęto węzeł H_PCONT_ST1 (Primary Containment Source Term 1). Węzeł ten reprezentuje hipotezy dotyczące zdarzenia inicjującego awarię w elektrowni jądrowej. Od wartości prawdopodobieństw a posteriori tego węzła zależą oceny ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii.

Analizę wrażliwości parametrów rozpatrywanej sieci bayesowskiej przeprowadzono następująco (rys.6.2): dla każdego przypadku, wyznaczono zbiór współczynników wrażliwości dla każdego parametru (wzór 4.4, str.36), następnie wyznaczono maksymalne wartości współczynników wrażliwości dla każdego parametru z wszystkich przypadków, w ostatnim etapie wyznaczono parametry, dla których wartości współczynnika wrażliwości są największe w obrębie rozpatrywanego węzła.

Procedura analizy wrażliwości parametrów składa się z następujących etapów:

- dla każdego przypadku tworzona jest macierz, zawierająca wyznaczone współczynniki wrażliwości (gdzie najpierw należało wyznaczyć współczynniki c_0, c_1, d_0, d_1 na podstawie wzoru 4.1, str.36), dla każdego parametru (rys.6.2-A). Każdy wiersz macierzy odpowiada jednemu parametrowi rozpatrywanej sieci bayesowskiej. Kolumny macierzy odnoszą się do poszczególnych hipotez węzła hipotez (w rozpatrywanym przypadku węzła H_PCONT_ST1). Każdy element macierzy odpowiada wykresowi podobnemu do wykresu 4.1 na str.37,



Rys. 6.2. Identyfikacja parametrów o maksymalnej wrażliwości

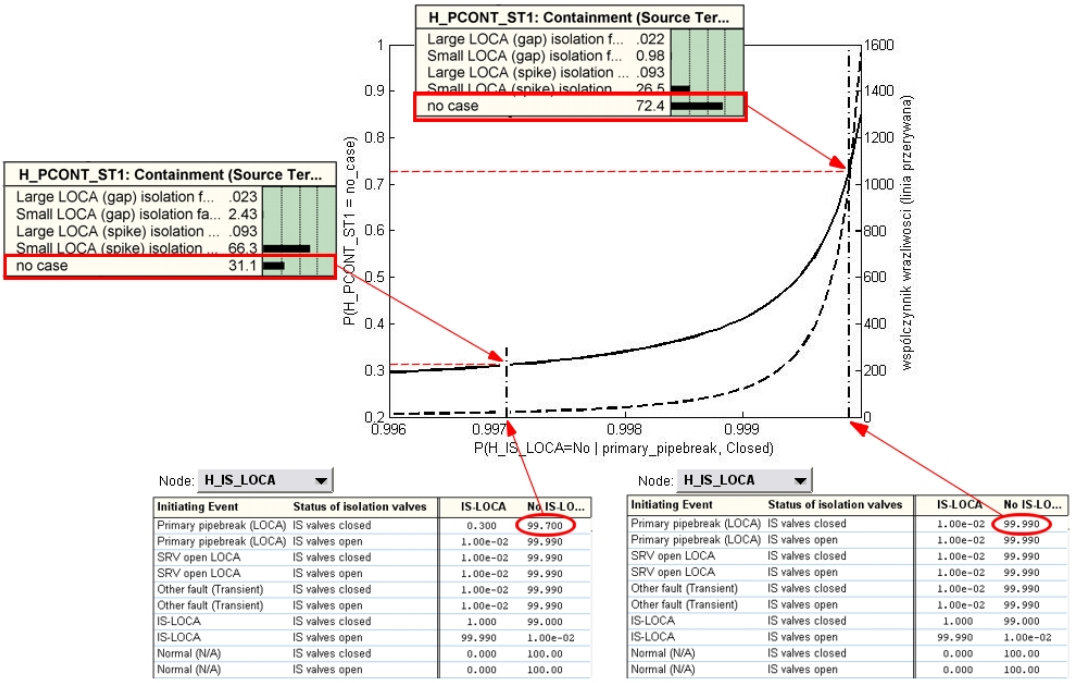
- następnie, wyznaczane są maksymalne wartości współczynników wrażliwości dla odpowiadających sobie elementów macierzy reprezentujących kolejne przypadki. Wyniki zapisywane są w nowej macierzy nazwanej *Maksymalne wartości dla wszystkich przypadków* (rys.6.2-B),
- w kolejnym etapie wyznaczane są parametry o największym współczynniku wrażliwości w obrębie każdego węzła, i stanowiące o wrażliwości całego węzła (rys.6.2-C).

Należy zwrócić uwagę na fakt, że w zaproponowanej procedurze wyznaczane są maksymalne wartości współczynników wrażliwości. Taka informacja, w przypadku tworzenia sieci bayesowskiej i ustalania prawdopodobieństw warunkowych, jest najbardziej potrzebna i oczekiwana, ze względu na przyjmowaną dokładność oszacowań tych prawdopodobieństw.

6.2.2. Interpretacja wyników

Analiza wrażliwości parametrów, przeprowadzona dla rozpatrywanej sieci bayesowskiej, pokazuje, że małe zmiany parametrów sieci mogą mieć różny (duży

lub mały) wpływ na wartości prawdopodobieństw a posteriori węzła hipotez. Na podstawie otrzymanych wyników można stwierdzić, że istnieje grupa kilkunastu parametrów (nazywanych słabymi punktami sieci - weak points of the network), których mała zmiana wartości powoduje duże zmiany prawdopodobieństw hipotez. Na rysunku 6.3 przedstawiono wykres reprezentujący zależność wartości prawdopodobieństwa hipotezy $H\_PCONT\_ST1 = no\_case$ od wartości parametru (prawdopodobieństwa warunkowego[1]) $H\_IS\_LOCA = No|primary\_pipebreak, Closed$. Dodatkowo linią przerywaną przedstawiono zależność wartości współczynnika wrażliwości $s1$ od wartości rozpatrywanego parametru. Analizując przedstawiony wykres, można zauważyć, że zmiana wartości parametru z 0,9999 na 0,997 powoduje zmiany prawdopodobieństwa hipotezy z 0,724 (interpretowanej jako raczej prawdziwa) na 0,311 (interpretowanej jako raczej fałszywa).

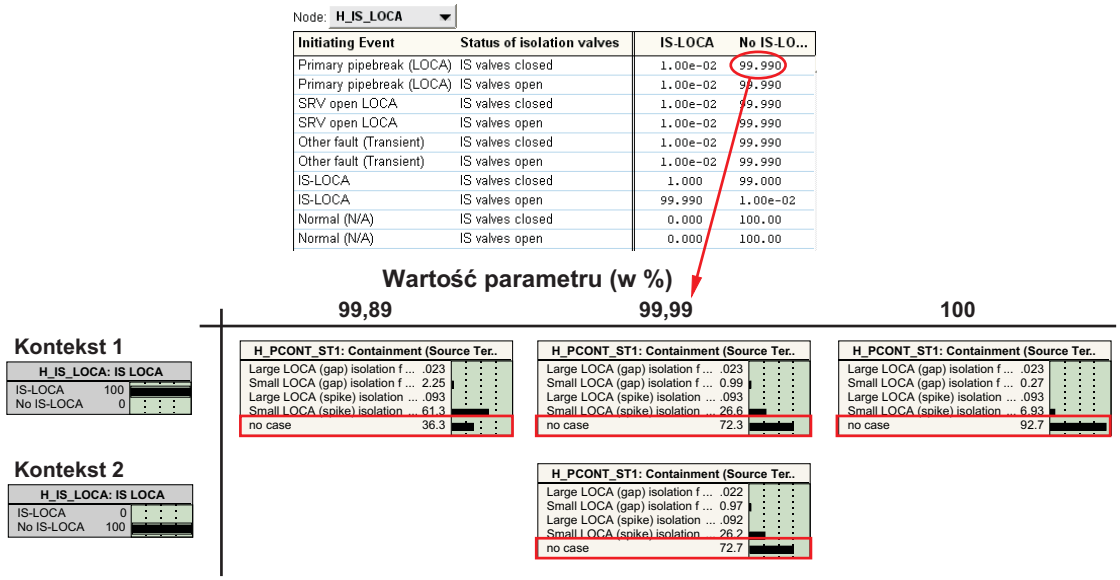


Rys. 6.3. Zależność wartości prawdopodobieństwa hipotezy $H\_PCONT\_ST1 = no\_case$ od wartości parametru $H\_IS\_LOCA = No|primary\_pipebreak, Closed$

Na rysunku 6.4 przedstawiono porównanie analizy wrażliwości węzłów obserwowanych i analizy wrażliwości parametrów dla węzła H_IS_LOCA. Z porównania tego wynika, że pomimo małej wrażliwości węzła hipotez na zmiany wartości węzła H_IS_LOCA (zmiana wartości z $IS - LOCA$ na $No\ IS - LOCA$ powoduje zmianę prawdopodobieństwa z 0,723 na 0,727) węzeł może posiadać parametry, których niewielka zmiana wartości wpływa na jakość wnioskowania (prawdopodobieństwo a posteriori hipotez).

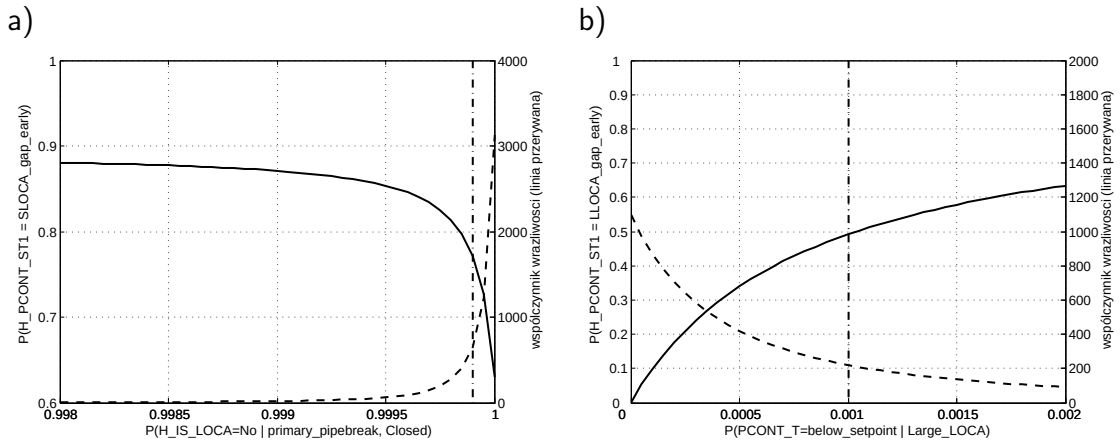
W trakcie badań stwierdzono, że współczynnik wrażliwości $s1$ wyznaczany jako $|\frac{d(P(h|e)(x))}{dx}|$ (wzór 4.4, str.36) nie dostarcza pełnej informacji o wrażliwości parametru.

[1]Ze względu na specyfikę programu NETICA, niektóre wartości prawdopodobieństw (w graficznych reprezentacjach fragmentów sieci bayesowskiej - węzłów, tablic prawdopodobieństw warunkowych) przedstawione są procentowo.



Rys. 6.4. Analiza wrażliwości węzłów obserwowanych i parametrów dla węzła H_IS_LOCA

Musi on być rozpatrywany wspólnie z przewidywanym zakresem zmian wartości prawdopodobieństwa hipotez, gdyż duża wartość parametru $s1$ nie gwarantuje dużej wrażliwości wartości prawdopodobieństwa hipotez na zmiany wartości parametru, ze względu na mały zakres zmian tych wartości. Sytuacja taka została przedstawiona na rys.6.5, gdzie w przypadku wykresu (rys.6.5a) widać, że współczynnik wrażliwości jest stosunkowo duży ($s1 = 648$), a zakres zmian wartości prawdopodobieństwa a posteriori hipotezy $P(H\_PCONT\_ST1 = SLOCA\_gap\_early)$ (dla zakresu zmiany wartości parametru 0,002) wynosi około 0,25. Natomiast w przypadku wykresu przedstawionego na rys.6.5b, gdzie wartość współczynnika wrażliwości jest dużo mniejsza ($s1 = 220$), zakres zmian wartości $P(H\_PCONT\_ST1 = LLOCA\_gap\_early)$ wynosi 0,63 (dla takiego samego zakresu wartości parametru).



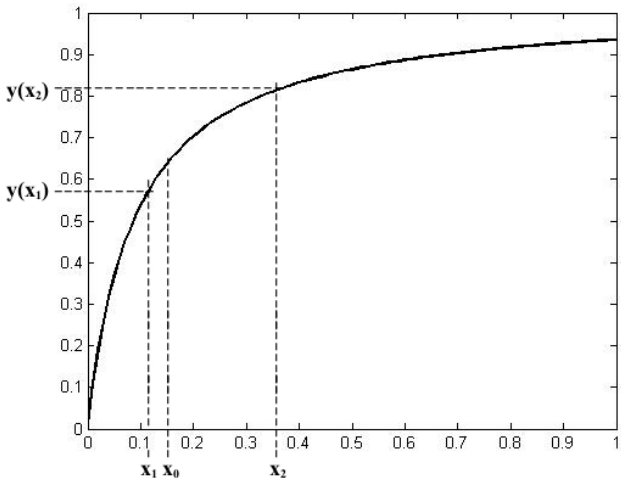
Rys. 6.5. Ilustracja sytuacji, w której współczynnik wrażliwości $s1$ nie dostarcza pełnej informacji o wrażliwości parametru: a) mały zakres zmian wartości prawdopodobieństwa hipotezy, b) duży zakres zmian wartości prawdopodobieństwa hipotezy

Dla uzyskania pełnej oceny wrażliwości wprowadzono dodatkową miarę wrażliwości $s2$,

reprezentującą zakres zmian wartości prawdopodobieństwa a posteriori hipotezy (6.1) dla założonego zakresu parametru q (6.2) (np. $q = 0,25$). Graficzna interpretacja współczynnika wrażliwości $s2$ została przedstawiona na rys.6.6. Współczynnik wrażliwości $s2$ dla przypadków przedstawionych na rysunku 6.5 wynosi odpowiednio $s2_a = 0,1378$ i $s2_b = 0,4604$ dla $q = 0,25$, co w właściwy sposób reprezentuje przedstawioną sytuację.

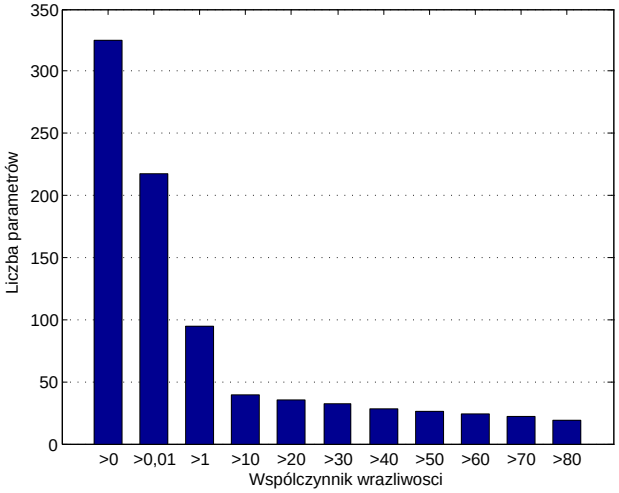
$$s2 = |y(x_1) - y(x_2)| \tag{6.1}$$

$$q = \frac{x_0 - x_1}{x_0} = \frac{x_2 - x_0}{1 - x_0} \tag{6.2}$$



Rys. 6.6. Graficzna interpretacja współczynnika wrażliwości $s2$ dla $q = 0,25$

Histogram przedstawiony na rysunku 6.7 przedstawia rozkład skumulowany maksymalnych wartości współczynnika wrażliwości $s1$, wyznaczonych dla rozpatrywanej sieci bayesowskiej. Obrazuje on liczbę parametrów wrażliwych i niewrażliwych.



Rys. 6.7. Histogram współczynnika wrażliwości $s1$

Pełne wyniki przeprowadzonej analizy wrażliwości parametrów rozpatrywanej sieci bayesowskiej znajdują się w [8].

6.3. Podsumowanie i wnioski

Przedstawiona została metoda analizy wrażliwości pozwalająca wyznaczyć słabe punkty sieci bayesowskiej, tzn. parametry, których zmiana wartości wywołuje duże zmiany wartości prawdopodobieństw a posteriori hipotez. Analiza taka została przeprowadzona dla sieci bayesowskiej do diagnozowania stanu reaktora w elektrowni jądrowej oraz oceny ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii. Rozpatrywana sieć bayesowska była opracowana przez specjalistów z dziedziny technologii reaktorów jądrowych. Ze względu na brak przykładów testujących, do badań postanowiono wygenerować zbiór 200 przykładów stanowiących konteksty (zbiory ustalonych węzłów obserwowanych) w analizie wrażliwości. Podczas badań stwierdzono, że współczynnik wrażliwości $s1$ w postaci pochodnej funkcji ma pewną wadę związaną z zakresem zmian wartości prawdopodobieństwa a posteriori hipotezy. Dlatego też, wprowadzono dodatkową miarę wrażliwości $s2$, reprezentującą wspomniany zakres.

Na podstawie otrzymanych wyników uzyskanych z przeprowadzonej analizy wrażliwości parametrów, można stwierdzić, że istnieje niewielka liczba parametrów mających duży wpływ na wartości prawdopodobieństw hipotez: dla współczynnika wrażliwości > 10 około 40 parametrów, dla współczynnika wrażliwości > 80 około 20 parametrów. Jednocześnie występuje relatywnie duża liczba parametrów nie mających wpływu (lub wpływających w małym stopniu) na wartości prawdopodobieństw hipotez.

Mała liczba parametrów wpływających w dużym stopniu na wartości węzła hipotez pozwala założyć, że zaproponowana metodologia jest racjonalna i wydajna.

Otrzymane wyniki przedstawionej analizy mogą służyć twórcom sieci w identyfikacji parametrów kluczowych dla procesu wnioskowania. Parametry te powinny zostać dokładnie zweryfikowane pod względem wpływu na jakość wnioskowania lub też, w niektórych przypadkach, sieć bayesowska powinna zostać zmodyfikowana w celu eliminacji wyznaczonych słabych punktów.

Rozdział 7

Podsumowanie i wnioski

W pracy przedstawiono wyniki badań dotyczące zastosowania analizy wrażliwości diagnostycznych sieci bayesowskich w celu doskonalenia tych sieci. W pracy opisano dwa rodzaje analizy wrażliwości: analizę węzłów obserwowanych oraz analizę parametrów sieci. W pierwszym przypadku zdefiniowana została wrażliwość węzłów obserwowanych. Natomiast w drugim przypadku szczegółowo omówiono metodę wyznaczania zaproponowanego współczynnika wrażliwości parametru oraz wprowadzono dodatkową miarę wrażliwości, reprezentującą zakres zmian wartości prawdopodobieństwa a posteriori hipotezy. Skuteczność tych metod w procesie doskonalenia diagnostycznych sieci bayesowskich wykazano na przykładach dwóch sieci. Pierwsza dotyczyła diagnozowania przekoszenia panwi łożyskowych i została zidentyfikowana w procesie uczenie. Druga została stworzona przez specjalistów z zakresu technologii reaktorów jądrowych i dotyczyła zagadnienia diagnozowania stanu reaktora w elektrowni jądrowej i oceny ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii. Dodatkowo, w pracy zamieszczono przegląd metod uczenia sieci bayesowskiej, zarówno w aspekcie jakościowym jak i w aspekcie ilościowym, oraz scharakteryzowano proces identyfikacji sieci bayesowskiej na podstawie wiedzy ekspertów z danej dziedziny. Oprócz tego przedstawiono również przegląd metod dyskretyzacji wartości cech ciągłych ze względu na konieczność odpowiedniego przygotowania danych do procesu uczenia sieci bayesowskiej. Analiza otrzymanych wyników badań pozwoliła na sformułowanie szeregu wniosków dotyczących zagadnienia doskonalenia sieci bayesowskich stosowanych w diagnostycznych systemach doradczych.

7.1. Wnioski ogólne

Na podstawie wyników otrzymanych w trakcie badań stwierdzono słuszność postawionych tez dotyczących optymalizacji parametrów i struktury sieci bayesowskich. Stwierdzono, że analiza wrażliwości węzłów obserwowanych diagnostycznej sieci bayesowskiej umożliwiła takie zmodyfikowanie struktury sieci, że jakość wnioskowania uległa znaczącej poprawie W przypadku sieci bayesowskiej dotyczącej diagnozowania stanu elektrowni jądrowej, na podstawie analizy wrażliwości wskazano parametry, których uważna weryfikacja, przez eksperta z dziedziny technologii reaktorów jądrowych, pod względem wpływu na jakość wnioskowania pozwoli na optymalizację tej sieci.

7.2. Wnioski szczegółowe

1. Wyniki weryfikacji sieci bayesowskich do diagnozowania przekoszenia panwi łożyskowych pokazują, że przemieszczenia podpór łożyskowych, mogą wywoływać skutki zbliżone do tych, które powstają przy przekoszeniu sąsiadujących z podpora panwi łożyskowych. Problem ten, można rozwiać, wprowadzając dodatkowe węzły reprezentujące cechy sygnału umożliwiające rozróżnienie tych niesprawności. Takie zadanie wymaga jednak dodatkowych badań.

2. W wielu przypadkach przeprowadzanie drugiego etapu optymalizacji struktury, jakim jest usunięcie węzłów, wyznaczonych w analizie wrażliwości węzłów obserwowanych nie jest konieczne. Jeżeli jakość otrzymanego, w wyniku optymalizacji modelu w postaci sieci bayesowskiej jest zadowalająca, a nie wykazuje on cech zbytniego dopasowania się do danych to drugi etap można pominąć. Usunięcie bowiem któregoś węzła, może przy pewnych warunkach (niekompletne, niepewne, błędne dane) prowadzić do sytuacji, w której niemożliwe będzie prowadzenie wiarygodnego procesu wnioskowania.

3. Ocena wrażliwości węzłów obserwowanych nie powinna być utożsamiana z oceną przydatności tych węzłów. Trzeba ją stosować ostrożnie, co nie zmienia faktu jej znaczącej, praktycznej przydatności.

4. Zaproponowany współczynnik wrażliwości $s1$, wyznaczany jako wartość bezwzględna pochodnej $\frac{d(P(h|e)(x))}{dx}$, musi być rozpatrywany wspólnie z przewidywanym zakresem zmian wartości prawdopodobieństwa hipotez. Wysoka wartość współczynnika wrażliwości $s1$ nie gwarantuje dużej wrażliwości wartości prawdopodobieństwa hipotez na zmiany wartości parametru, ze względu na mały zakres zmian tych wartości. Dla uzyskania pełnej oceny wrażliwości wprowadzono dodatkową miarę wrażliwości $s2$, reprezentującą zakres zmian wartości prawdopodobieństwa a posteriori hipotezy dla założonego zakresu parametru q.

5. Wyjątkowo przydatną klasą modeli są modele formułowane w postaci sieci bayesowskich. Rozpatrywane sieci bayesowskie dotyczące diagnozowania przekoszeń panwi łożyskowych oraz diagnozowania stanu reaktora jądrowego mogą być wykorzystane w modułach wnioskujących systemu doradczego K015 [58, 98]. System doradczy K015 jest rekonfigurowalnym systemem hybrydowym, przeznaczonym do oceniania stanu technicznego różnych obiektów, na podstawie dostępnych cech sygnałów diagnostycznych. Głównymi elementami modułów wnioskujących, a zarazem całego systemu doradczego K015, są procedury wnioskujące, wyznaczone na podstawie modeli diagnostycznych. Zastosowanie sieci bayesowskich w diagnostycznych systemach doradczych, będące obecnie jednym z istotniejszych zastosowań sztucznej inteligencji, daje szerokie korzyści i stwarza interesujące perspektywy w rozwoju narzędzi wspomagających podejmowanie decyzji.

7.3. Kierunki dalszych badań

Przedstawione analizy wrażliwości wykorzystane do doskonalenia diagnostycznych sieci bayesowskich mogą znaleźć szerokie zastosowanie w procesie identyfikacji tych sieci zarówno na podstawie danych jak i wiedzy ekspertów. Rozpatrywane sieci bayesowskie wymagają testowania w warunkach rzeczywistych co związane jest z dalszymi badaniami. Ponadto, interesującą wydaje się możliwość zastosowania rozpatrywanych sieci w systemie K015.

Szczególnie poznawczą wydaje się być próba analizy wrażliwości N-tego rzędu (badanie wpływu zmian kombinacji wartości węzłów obserwowanych lub/i parametrów na prawdopodobieństwa a posteriori hipotez). Jest to zadanie NP-trudne, dlatego też przeprowadzenie tej analizy wymaga zastosowania metod heurystycznych, np. algorytmów ewolucyjnych. Rozwojowym obszarem badań może być również wykorzystanie algorytmów ewolucyjnych w procesie uczenia struktury sieci bayesowskiej na podstawie danych, gdzie podobnie jak w przypadku analizy wrażliwości N-tego rzędu mamy do czynienia z zadaniem NP-trudnym.

Literatura

[1] Abramson B.: *The design of Belief Network-based Systems for Price Forecasting*. Computers and Electrical Engineering, 20(2), s. 163–180, 1994.

[2] Acid S., de Campos, L. M.: *BENEDICT: An algorithm for learning probabilistic belief networks*. Proceedings of the 6th International Conference IPMU'96, s. 979–984, Granada, Spain, 1996.

[3] Acid S., de Campos, L. M., Fernández-Luna J. M., Rodriguez S., Rodriguez J. M., Salcedo J. L.: *A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service*. Artificial Intelligence in Medicine, 30(3), s. 215–232, 2004.

[4] Akaike H.: *A new look at the statistical model identification*. IEEE Transactions on automatic control, 19(6), s. 716–723, 1974.

[5] Almond R.: *Graphical Belief Modelling*. Chapman & Hall, 1995.

[6] Banaszek S.: *Obliczenia pęknięć wirnika turbozespołu 13K215*. IMP PAN, Opracowanie wewnętrzne, Gdańsk, 2004.

[7] Bayes Net Toolbox for Matlab. `http://bnt.sourceforge.net/`.

[8] Bednarski M.: *Metody doskonalenia sieci bayesowskich stosowanych w diagnostycznych systemach doradczych*. Rozprawa doktorska, Politechnika Śląska, Wydział Mechaniczny Technologiczny, Gliwice, Lipiec 2006.

[9] Bednarski M., Cholewa W., Frid W.: *Identification of sensitivities in Bayesian networks*. Engineering Applications of Artificial Intelligence, 17(4), s. 327–335, 2004.

[10] Belief Network (BN) PowerConstructor. `http://www.cs.ualberta.ca/~jcheng/bnpc.htm`.

[11] BNT Structure Learning Package. `http://banquiseasi.insa-rouen.fr/projects/bnt-slp/`.

[12] Bouckaert R. R.: *Probabilistic Network Construction Using the Minimum Description Length Principle*. Symbolic and Quantitative Approaches to Reasoning and Uncertainty, European Conference, ECSQARU'93, 747 of *Lecture Notes in Computer Science*, s. 41–48. Springer, 1993.

[13] Breese J., Horvitz E., Peot M., Gay R., Quentin G.: *Automated decisionanalytic diagnosis of thermal performance in gas turbines*. The International Gas Tur-

bine and Aeroengine Congress and exposistion. American Society of Mechanical Engineers, 1992.

[14] Bronsztejn I. N., Siemiendiajew K. A.: *Matematyka. Poradnik encyklopedyczny*. Wydawnictwo Naukowe PWN, Warszawa, 1995.

[15] Buntine W.: *Theory Refinement on Bayesian Networks*. Proceedings of the 7th Annual Conference on Uncertainty in Artificial Intelligence (UAI-91), s. 52–60. Morgan Kaufmann, 1991.

[16] de Campos, L. M., Huete J. F.: *A new approach for learning belief networks using independence criteria*. International Journal of Approximate Reasoning, 24(1), s. 11–37, April 2000.

[17] Castillo E., Gutiérrez J. M., Hadi A. S.: *Expert Systems and Probabilistic Network Models*. Springer, 1997.

[18] Catlett, J.: *On changing continuous attributes into ordered discrete attributes*. Proceedings of the European working session on learning on Machine learning, s. 164–178. Springer, 1991.

[19] Cempel C.: *Podstawy wibroakustycznej diagnostyki maszyn*. WNT, Warszawa, 1982.

[20] Cempel C.: *Diagnostyka wibroakustyczna maszyn*. PWN, Warszawa, 1989.

[21] Cerquides J., Màntaras R. L.: *Proposal and Empirical Comparison of a Parallelizable Distance-Based Discretization Method*. Proceedings of the Third International Conference on Knowledge Discovery and Data Mining, s. 139–142, 1997.

[22] Chan C. C., Batur C., Srinivasan A.: *Determination of quantization intervals in rule based model for dynamic systems*. Proceedings of the IEEE Conference on Systems and Man and and Cybernetics, s. 1719–1723, 1991.

[23] Cheng J., Bell D. A., Liu W.: *An algorithm for Bayesian belief network construction from data*. Proceedings of AI & STAT'97, s. 83–90, 1997.

[24] Cheng J., Greiner R., Kelly J., D. B., Liu W.: *Learning Bayesian networks from data: an information-theory based approach*. Artificial Intelligence, 137(1-2), s. 43–90, 2002.

[25] Chickering D. M.: *Optimal Structure Identification With Greedy Search*. Journal of Machine Learning Research, 3(Nov), s. 507–554, 2002.

[26] Cholewa W.: *e-Tutor*. Methods of Artificial Intelligence, AI-METH Series, Gliwice, November 2002.

[27] Cholewa W., Kaźmierczak J.: *Diagnostyka Techniczna Maszyn. Przetwarzanie cech sygnałów*. Wydawnictwo Politechniki Śląskiej, Gliwice, 1995.

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

[29] Cholewa W., Moczulski W.: *Diagnostyka techniczna maszyn. Pomiary i analiza sygnałów*. Wydawnictwo Politechniki Śląskiej, Gliwice, 1993.

[30] Cholewa W., Pedrycz W.: *Systemy doradcze*. Politechnika Śląska, Gliwice, 1987.

[31] Chow C. K., Liu C. N.: *Approximating discrete probability distributions with dependence trees*. IEEE Transactions on Information Theory, 14(3), s. 462–467, 1968.

[32] Cooper G. F., Herskovits E.: *A Bayesian Method for the Induction of Probabilistic Networks from Data*. Machine Learning, 9(4), s. 309–347, October 1992.

[33] Coupé V. M. H., van der Gaag, L. C.: *Properties of Sensitivity Analysis of Bayesian Belief Networks*. Annals of Mathematics and Artificial Intelligence, 36(4), s. 323–356, 2002.

[34] Coupé V. M. H., van der Gaag, L. C., Habbema J.: *Sensitivity analysis: an aid for belief-network quantification*. Knowledge Engineering Review, 15(3), s. 1–18, 2000.

[35] Coupé V. M. H., Jensen F. V., Kjærulff U., van der Gaag, L. C.: *A Computational Architecture for N-way Sensitivity Analysis of Bayesian Networks*. Aalborg University, 2000.

[36] Cowell R. G., Dawid A. P., Lauritzen S. L.: *Probabilistic Networks and Expert Systems*. Springer, 1999.

[37] Dougherty J., Kohavi R., Sahami M.: *Supervised and Unsupervised Discretization of Continuous Features*. Proceeding of the 12th International Conference on Machine Learning, s. 194–202, 1995.

[38] DT200-1 - Systemem diagnostyki turbozespołów energetycznych. `https://www.kpkm.polsl.pl/PROJEKTY/SystemDT200-1/default.htm`.

[39] Dubik H. i inni: *Containment protection during severe accidents*. Proceedings of SMIRT 16, Washington, USA, August 2001.

[40] Fayyad U. M., Irani K. B.: *Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning*. Proceeding of the 13th International Joint Conference on Artificial Inteligence, s. 1022–1027. Morgan Kaufmann, 1993.

[41] Frid W., Knochenhauer M., Bednarski M.: *Development of a Bayesian belief network for a Boiling Water Reactor during fault conditions*. Computer Assisted Mechanics and Engineering Sciences, 12(2-3), s. 133–145, 2005.

[42] Friedman N.: *The Bayesian Structural EM Algorithm*. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98), s. 129–138. Morgan Kaufmann, 1998.

[43] Friedman N., Goldszmidt M.: *Discretizing Continuous Attributes While Learning Bayesian Networks*. Fifteenth International Conference on Machine Learning (ICML), s. 157–165, 1996.

[44] GeNIe & SMILE. `http://genie.sis.pitt.edu/`.

[45] Goldman R.: *A Probabilistic Approach to Language Understanding*. Department of Computer Science, Brown University. Technical Report CS-90-34, 1990.

[46] Grindon E. i inni: *A Rapid Response Source Term Indicator Based on Plant Status for Use in Emergency Response (STERPS)*. Proceedings of FISA-2003, EU research in reactor safety, November 10-13, 2003.

[47] Guérin-Dugué A. i inni: *Deliverable R3-B4-P - Task B4: Benchmarks*. Elena-NervesII Ęnhanced Learning for Evolutive Neural Architecture", ESPRIT-Basic Research Project Number 6891, June 1995. `ftp://ftp.dice.ucl.ac.be/pub/neural-nets/ELENA/Benchmarks.ps.Z`.

[48] Heckerman D.: *A Tutorial on Learning Bayesian Networks*. Microsoft Corporation, MSR-TR-95-06, 1995.

[49] Heckerman D., Geiger D., Chickering D. M.: *Learning Bayesian Networks: The Combination of Knowledge and Statistical Data*. Machine Learning, 20(3), s. 197–243, September 1995.

[50] Herskovits E., Cooper G. F.: *Kutató: An Entropy-Driven System for Construction of Probabilistic Expert Systems from Databases*. Proceedings of the 6th Annual Conference on Uncertainty in Artificial Intelligence (UAI-91), s. 54–62. Elsevier Science Publishing Company, Inc., 1990.

[51] Ho K. M., Scott P. D.: *Zeta: A Global Method for Discretization of Continuous Variables*. Proceedings of Third International Conference on Knowledge Discovery and Data Mining (KDD-97), s. 191–194, 1997.

[52] Holte R. C.: *Very Simple Classification Rules Perform Well on Most Commonly Used Datasets*. Machine Learning, 11(1), s. 63–91, 1993.

[53] Hugin Expert. `http://www.hugin.com/`.

[54] Jensen F. V.: *An Introduction to Bayesian Networks*. Springer, 1996.

[55] Jensen F. V.: *Bayesian Networks and Decision Graphs*. Springer, 2001.

[56] Johansson M.: *Input data for the project STERPS*. OKG report 2002-02102, February 2002.

[57] Kerber R.: *ChiMerge: Discretization of Numeric Attributes*. Proceedings of the 10th National Conference on Artificial Intelligence (AAAI-92), s. 123–128. AAAI Press and MIT Press, 1992.

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

[59] Kiciński J., Prońska A.: *Identyfikacja modelu obliczeniowego turbozespołu 13K215*. IMP PAN, Opracowanie wewnętrzne 4068/04, Gdańsk, 2004.

[60] Kjærulff U., van der Gaag, L. C.: *Making Sensitivity Analysis Computationally Efficient*. Proceedings of the 16th Annual Conference on Uncertainty in Artificial Intelligence (UAI-00), s. 317–325, San Francisco, CA, 2000. Morgan Kaufmann.

[61] Korbicz J., Kościelny J., Kowalczuk Z., Cholewa W. (red.): *Diagnostyka procesów. Modele, Metody Sztucznej Inteligencji, Zastosowania*. WNT, 2002.

[62] Lam W., Bacchus F.: *Learning Bayesian Belief Networks: An Approach Based on the MDL Principle*. Computational Intelligence, 10(3), s. 269–293, 1994.

[63] Laskey K. B.: *Sensitivity analysis for probability assessments in Bayesian networks*. IEEE Transactions on Systems, Man, and Cybernetics, 25(6), s. 901–909, 1995.

[64] Lauritzen S. L.: *Grpahical Models*. Oxford University Press, 1996.

[65] Leray P., Francois O.: *BNT Structure Learning Package: Documentation and Experiments*. Laboratoire PSI, November 2004.

[66] Liu H., Hussain F., Tan C. L., Dash M.: *Discretization: An Enabling Technique*. Data Mining and Knowledge Discovery, 6(4), s. 393–423, 2002.

[67] Liu H., Setiono R.: *Chi2: Feature selection and discretization of numeric attributes*. Proceedings of Seventh IEEE International Conference on Tools with Artificial Intelligence, s. 388–391. IEEE Computer Society, 1995.

[68] Łączkowski, R.: *Drgania elementów turbin cieplnych*. WNT, Warszawa, 1974.

[69] Łączkowski, R.: *Wibroakustyka maszyn i urządzeń*. WNT, Warszawa, 1983.

[70] Łuczak, M.: *Badania symulacyjne wpływu defektu w postaci przekoszenia panwi na własności dynamiczne turbozespołu 200MW*. IMP PAN, Opracowanie wewnętrzne, Gdańsk, 2004.

[71] Mántaras R. L.: *A Distance-Based Attribute Selection Measure for Decision Tree Induction*. Machine Learning, 6(1), s. 81–92, 1991.

[72] Michalski R. (red.): *Diagnostyka maszyn roboczych. Detekcja, relacje, wnioskowanie hybrydowe*. Instytut Technologii Eksploatacji, Radom, 2004.

[73] Microsoft Visual C++ Developer Center. `http://msdn.microsoft.com/visualc/`.

[74] Mitchell J. S.: *Introduction to Machinery Analysis and Monitoring*. Pennwell Books, 1993.

[75] Mitchell T.: *Machine Learning*. McGraw Hill, 1997.

[76] Moczulski W.: *Metody pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*. Zeszyty Naukowe Politechniki Śląskiej. Wydawnictwo Politechniki Śląskiej, Gliwice, 1997.

[77] Moczulski W.: *Pozyskiwanie wiedzy deklaratywnej i proceduralnej*. Methods of Artificial Intelligence in Mechanics and Mechanical Engineering, AI-MECH 2000, s. 55–74, Gliwice, 14-17 listopad 2000.

[78] Moczulski W.: *Diagnostyka Techniczna. Metody pozyskiwania wiedzy*. Wydawnictwo Politechniki Śląskiej, Gliwice, 2002.

[79] Morel J.: *Drgania maszyn i diagnostyka ich stanu technicznego*. PTDT, Warszawa, 1994.

[80] MSBNx - Bayesian Network Editor and Tool Kit. `http://research.microsoft.com/adapt/MSBNx/`.

[81] Nadi F., Agogino A., Hodges D.: *Use of Influence Diagrams and Neural Networks in Modeling Semiconductor Manufacturing Processes*. IEEE Transactions on Semiconductor Manufacturing, 4(1), 1991.

[82] Neapolitan R. E.: *Learning Bayesian Networks*. Pearson Prentice Hall, 2004.

[83] Netica API Programmer's Library C Language Version, Reference Manual Version 2.15. `http://www.norsys.com/onLineAPIManual/index.html`.

[84] Niziński S., Michalski R.: *Diagnostyka obiektów technicznych*. Instytut Technologii Eksploatacji, Warszawa-Sulejówek-Olsztyn-Radom, 2002.

[85] Norsys Software Corp. `http://www.norsys.com`.

[86] Orłowski Z.: *Diagnostyka w życiu turbin parowych*. WNT, Warszawa, 2001.

[87] Peña J. M., Lozano J. A., Larrañga P.: *An improved Bayesian structural EM algorithm for learning Bayesian networks for clustering*. Pattern Recognition Letters, 21(9), s. 779–786, July 2000.

[88] Pearl J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[89] Perner P., Trautzsch S.: *Multi-Interval Discretization Methods for Decision Tree Learning*. Advances in Pattern Recognition, Lecture Notes in Computer Science, 1451, s. 475–482, 1998.

[90] Rabaséda S., Rakotomalala R., Sebban M.: *A Comparison of Some Contextual Discretization Methods*. Informatics and Computer Science, 92(1-4), s. 137–157, 1996.

[91] Rybczyński J.: *Przemieszczenia termiczne i montażowe podpór łożyskowych*. IMP PAN, Opracowanie wewnętrzne 4155/04, Gdańsk, 2004.

[92] Schwarz G.: *Estimating the Dimension of a Model*. The Annals of Statistics, 6(2), s. 461–464, 1978.

[93] Singh M., Valtorta M.: *Construction of Bayesian network structures from data: A brief survey and an efficient algorithm*. International Journal of Approximate Reasoning, 12(2), s. 111–131, 1995.

[94] Sobczak W., Malina W.: *Metody selekcji informacji*. WNT, 1978.

[95] Sobczak W., Malina W.: *Metody selekcji i redukcji informacji*. WNT, 1985.

[96] Spiegelhalter D., Franklin R., Bull K.: *Assessment Criticism and Improvement of Imprecise Subjective Probabilities for a Medical Expert System*. The Fifth Workshop on Uncertainty in Artificial Intelligence, 1989.

[97] Spirtes P., Glymour C., Scheines R.: *Causation, Prediction, and Search*. The MIT Press, 2001.

[98] System doradczy K015. `https://www.kpkm.polsl.pl/k015/`.

[99] The MathWorks, Inc. `http://www.mathworks.com/`.

[100] Wang X. H., Zheng B., Good W. F., King J. L., Chang Y. H.: *Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network*. International Journal of Medical Informatics, 54(2), 1999.

[101] Woodberry O., Nicholson A. E., Korb K. B., Pollino C.: *Parameterising Bayesian Networks*. Australian Conference on Artificial Intelligence, Lecture Notes in Computer Science, s. 1101–1107. Springer, 2004.

[102] Żółtowski, B.: *Podstawy diagnostyki maszyn*. Wydawnictwo Akademii Techniczno-Rolniczej, Bydgoszcz, 1996.

Streszczenie

Praca poświęcona jest metodologii doskonalenia diagnostycznych sieci bayesowskich na podstawie dwóch rodzajów analiz wrażliwości: analizy węzłów obserwowanych i analizy parametrów sieci. W pierwszym przypadku zdefiniowana została wrażliwość węzłów obserwowanych. Natomiast, w drugim przypadku szczegółowo omówiono metodę wyznaczania zaproponowanego współczynnika wrażliwości parametru oraz wprowadzono dodatkową miarę wrażliwości, reprezentującą zakres zmian wartości prawdopodobieństwa a posteriori hipotezy.

W pracy przedstawiono koncepcję sieci bayesowskiej wraz z twierdzeniem Bayesa, na którym została oparta. Dodatkowo zwrócono uwagę na właściwość sieci bayesowskiej jaką jest warunkowa niezależność węzłów. Zamieszczono przegląd metod uczenia sieci bayesowskiej, zarówno w aspekcie jakościowym jak i w aspekcie ilościowym, oraz scharakteryzowano proces identyfikacji sieci bayesowskiej na podstawie wiedzy ekspertów z danej dziedziny.

Zaprezentowano wyniki badań nad wykorzystaniem analizy wrażliwości węzłów obserwowanych do optymalizacji struktury sieci bayesowskiey na podstawie kryterium jakości wnioskowania. Przedmiotem badań, w tym przypadku, była sieć bayesowska uzyskana w procesie uczenia sieci na podstawie danych dotyczących przekoszenia panwi łożyskowych oraz przemieszczenia podpór łożyskowych.

Przedstawiono analizę wrażliwości parametrów sieci bayesowskiej, na podstawie której, jak wykazano, możliwe jest przeprowadzenie procesu doskonalenia tej sieci. Badana sieć bayesowska została stworzona przez specjalistów z dziedziny technologii reaktorów jądrowych i dotyczyła zagadnienia diagnozowania stanu reaktora w elektrowni jądrowej oraz oceny ilości i składu produktów radioaktywnych (Source Terms) uwolnionych do atmosfery w wyniku awarii.

Summary

The thesis is devoted to methodology of Bayesian network improvement on the basis of two kinds sensitivity analysis: sensitivity analysis of observables and sensitivity analysis of network parameters. In the first case sensitivity of observables was defined. Whereas, in the other case, the method of parameter sensitivity coefficient determination was discussed in detail and additionally measure of sensitivity, representing the range of the a posteriori hypothesis probability value changes, was discussed.

In the thesis the concept of Bayesian network along with Bayes theorem, on which it is based, was presented. The additional attention was bring to conditional independence of nodes as a property of Bayesian network. Survey of learning Bayesian network methods, both in qualitative aspect and in quantitative aspect, was placed and the process of Bayesian network identification on the basis of expert knowledge from given domain, was characterized.

The results of research on sensitivity analysis of observables used to Bayesian network structure optimization on the basis of inferring quality criterion, was presented. In this case, the object of research was the Bayesian network obtained in the process of network learning on the data, which cover angle misalignment and bearing displacement.

The possibility of conducting the network improvement process, on the basis of the sensitivity analysis of Bayesian network parameters was demonstrated. The considered network was made by specialists from domain of nuclear reactors technology and it concerned the matter of diagnosis reactor state in nuclear plant and estimation of Source Terms (the quantity, characteristics and timing of the release of radioactivity to the environment) in case of accident.