

dr hab. inż. Rafał Stanisławski, prof. uczelni
Katedra Informatyki
Wydział Elektrotechniki, Automatyki i Informatyki
Politechnika Opolska
e-mail: r.stanislawski@po.edu.pl

Opole, 2021-04-23,

S E K R E T A R I A T
Rady Dyscypliny AEE

Wpłynęło dnia 7.05.2021
Zarejestrowano pod nr
Podpis *dm*

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy:

Statistical reasoning analysis of fault occurrences in industrial applications

Autor rozprawy: mgr inż. Tian Cong

Promotor rozprawy: dr hab. inż. Jerzy Baranowski

Promotor pomocniczy: dr James R. Ottewill

Niniejsza recenzja została opracowana na zlecenie Przewodniczącego Rady Naukowej Automatyka, Elektronika i Elektrotechnika Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie dra hab. inż. Ryszarda Sroki, prof. uczelni, z 15 lutego 2021 roku.

1. Ocena formalna pracy

Przedłożona do oceny rozprawa doktorska zawiera łącznie 113 stron podzielonych na siedem rozdziałów, literatury liczącej 238 pozycji, oraz dwóch dodatków zawierających definicje terminów stosowanych w systemach monitorowania stanu procesu i listę publikacji współautorskich Doktorantki. Ponadto w pracy zawarto streszczenia w języku polskim i angielskim. Dysertacja została napisana w języku angielskim.

W rozdziale pierwszym przedstawiono wprowadzenie do problematyki poruszanej w pracy. Przedstawiono sformułowanie problemu badawczego, scharakteryzowano cel pracy, oraz określono szczegółowe zadania do zrealizowania w obszarze monitorowania stanu procesów (ang. *Process Condition Monitoring* – PCM) wraz z uzasadnieniem ich podjęcia. Ponadto w rozdziale pierwszym omówiono strukturę przedmiotowej dysertacji.

W rozdziale drugim przedstawiono wprowadzenie do problematyki PCM. Przedstawiono podstawowe zadania stawiane przed systemami PCM oraz scharakteryzowano poszczególne procesy związane z tymi systemami. Przedstawiono podstawową klasyfikację narzędzi modelowania/identyfikacji systemów, omówiono metody stosowane w monitoringu procesów jednomodowych (ang. *single mode process*) i wielomodowych (ang. *multimode process*) oraz opisano procesy podejmowania decyzji (ang. *decision-making process*). W rozdziale scharakteryzowano również możliwości systemów PCM pracujących w środowisku przemysłowym oraz problemy i wyzwania stojące przed tymi systemami.

Rozdział trzeci został poświęcony autorskiemu algorytmowi detekcji uszkodzeń opartemu na klasyfikatorze binarnym, tzw. BaFFle (ang. *Binary Classifier for Fault Detection*). W rozdziale, w pierwszej kolejności omówiono możliwości adaptacji różnych algorytmów detekcji uszkodzeń. Przedstawiono jednowymiarową i wielowymiarową metodę wykresu kontrolnego Shewharta (ang.

P. J.

Shewhart control chart) oraz zaproponowano narzędzia implementujące tą metodę w zadaniu monitorowania procesów wielowymiarowych. Następnie poddano analizie metodę analizy głównych składowych (ang. *Principal Component Analysis – PCA*) oraz dokonano implementacji wybranych wielowymiarowych estymatorów rozkładu gęstości prawdopodobieństwa. W oparciu o ww. elementy Autorka opracowała algorytm BaFFle, który jest rekomendowany dla procesów jednomodowych.

Wprowadzenie procesu Dirichleta opartego na modelu mieszanin gaussowskich (ang. *Dirichlet process – Gaussian Mixture Models– DP-GMM*) oraz jego implementacja w zadaniu klasteryzacji zostały zrealizowane w rozdziale czwartym. W pierwszej kolejności przedstawiono wybrane rozkłady prawdopodobieństwa tj. zero-jedynkowy rozkład Bernoulliego, rozkład dwumianowy, wielomianowy, Beta i uogólniony rozkład Dirichleta, oraz przedstawiono zależności pomiędzy wybranymi rozkładami w kontekście twierdzenia Bayesa. Następnie przedstawiono różne wersje modelu mieszanin gaussowskich (ang. *Gaussian Mixture Model – GMM*) obejmujących m. in. skończone i nieskończone implementacje modelu, jak również model Dirichleta. Ponadto dokonano analizy zależności rozkładu DP-GMM z odwrotnym rozkładem Wisharda (ang. *inverse-Wishart distribution*). W rozdziale zawarto również analizę metod implementacji obliczeniowej skończonej i nieskończonej wersji modelu GMM oraz modelu DP-GMM. W końcowej części rozdziału poddano szczegółowej analizie problematykę doboru parametrów modelu DP-GMM oraz ich wpływu na dokładność modelowania. Należy zaznaczyć, że rozważania teoretyczne zostały dokładnie poparte badaniami symulacyjnymi. Rozdział czwarty został zwieńczony prezentacją implementacji modeli w budowanym w ramach pracy narzędziu diagnostycznym.

Rozdział piąty został poświęcony detekcji anomalii w systemach wielomodowych. Do rozróżniania stanów procesów wielomodowych i detekcji anomalii Kandydatka wykorzystuje specyficzne rozszerzenie filtru Kalmana pozwalające na estymację stanów, parametrów i zakłóceń procesu, który został nazwany FKF (ang. *Field Kalman Filer*). Filtr FKF został zaimplementowany do wielowymiarowego, autoregresyjnego modelu w postaci równań stanu (ang. *Multivariate Autoregressive State-Space – MARSS*). Ostatecznie w rozdziale został opracowany algorytm implementacji ww. narzędzia w budowanym systemie diagnostycznym oraz dokonane zostają obszernie badania symulacyjne wprowadzonej metody.

Rozdział szósty pracy ma charakter praktyczny. Badania systemu diagnostycznego opracowanego w ramach dysertacji przeprowadzono na ogólnodostępnym zestawie danych o nazwie PRONTO, który opiera się na pomiarach dokonanych w wielofazowej instalacji przepływowej. Zestaw obejmuje dane z heterogenicznych źródeł, obejmujących pomiary sygnałów procesowych (ultradźwiękowe pomiary przepływu i ciśnienia wysokiej częstotliwości), zapisy alarmów, dzienniki operacji i nagrania wideo. W zestawie dostępne są dane pobrane w różnych warunkach operacyjnych, obejmujących również stany awaryjne. W omawianym rozdziale poddano analizie metody oparte na klasyfikatorze binarnym (rozpatrywanym w rozdziale trzecim) oraz metody implementujące filtr Kalmana (rozdział piąty). Rozdział został zwieńczony analizą efektywności poszczególnych metod.

W rozdziale siódmym zostało zawarte podsumowanie oraz wytyczono planowane kierunki dalszych prac.

Układ pracy jest poprawny. Treści dysertacji zostały logicznie podzielone na poszczególne rozdziały. Jedyną wątpliwość budzi kolejność dwóch pierwszych rozdziałów. Pierwej kolejności w rozdziale pierwszym przedstawiono opis kontrybucji pracy, a w drugim zaprezentowano wprowadzenie do problematyki PCM. Odwrócenie kolejności rozdziałów byłoby, zdaniem recenzenta, bardziej

R. Gh

logiczne, gdyż w obecnej strukturze w pierwszej kolejności czytelnik zapoznaje się z celami pracy, a tło w postaci problematyki PCM poznaje później. Układ poszczególnych rozdziałów nie budzi wątpliwości. Ponadto każdy rozdział merytoryczny jest zakończony wnioskami, co dodatkowo ułatwia czytanie dysertacji.

Warta szczególnie podkreślenia jest bardzo dobra strona edycyjna pracy. Rozdziały, podrozdziały, sekcje, nagłówki, stopki itp. utrzymane są tej samej konwencji, przez co praca bardzo dobrze wygląda. Rysunki są bardzo starannie przygotowane z dbałością o estetyczne i precyzyjne przedstawienie graficzne wyników badań oraz schematów blokowych. Język użyty w pracy jest czytelny i precyzyjny, przez co pracę czyta się dość dobrze. Jediną wątpliwość budzi w tym aspekcie używanie przez Doktorantkę bardzo dużej ilości skrótów, które nie zostały w żadnym miejscu razem zestawione. Rozprawa jest również bardzo dobrze przygotowana redakcyjnie, nieliczne błędy typograficzne zostały wypunktowane w dalszej części recenzji.

2. Ocena merytoryczna pracy

Wykrywanie usterek/anomalii i diagnostyka w procesach przemysłowych od długiego czasu stanowi jeden z ważniejszych aspektów poprawy niezawodności systemów sterowania i regulacji automatycznej. W ciągu ostatnich kilkudziesięciu lat odnotowano w tym obszarze bardzo dużą liczbę prac prezentujących różnorodne podejścia do tej problematyki. Jedne z bardziej popularnych i skutecznych metodologii stosowanych w tym obszarze są rozwiązania oparte na obserwatorach procesów, systemy oparte na narzędziach identyfikacji oraz podejścia statystyczne wykorzystujące twierdzenie Bayesa. Spośród ww. metod podejścia oparte na obserwatorach są zwykle skutecznie stosowane w połączeniu z dynamicznymi systemami stochastycznymi opisanymi w przestrzeni stanów. Większość metod diagnostycznych dla systemów stochastycznych rozważa modele z zakłóceniami gaussowskimi, w których są stosowane filtry Kalmana. Jednak w przypadku wielu procesów przemysłowych bardziej skutecznym podejściem okazują się rozwiązania oparte na rozkładach gęstości prawdopodobieństwa sygnałów stochastycznych, zamiast ich wartości chwilowych. W tym kontekście efektywne wykorzystanie informacji statystycznych stało się podejściem dającym nowe możliwości. Doprowadziło to do intensywnego rozwoju tej klasy metod w ciągu ostatnich kilkunastu lat. W związku z powyższym Autorka podejmuje pracę w ważnym i aktualnym obszarze badań, który daje możliwości osiągnięcia istotnych wyników zarówno w ujęciu teoretycznym, jak również praktycznym. Należy również zaznaczyć, że tematyka pracy jest jednoznacznie umiejscowiona w dyscyplinie Automatyka, Elektronika i Elektrotechnika w dziedzinie nauk inżynieryjno-technicznych.

Autorka rozprawy postawiła sobie ambitny i ważny cel zbudowania systemu pozwalającego na monitorowanie procesów o różnym stopniu złożoności. Osiągnięcie celu wymagało opracowania odrębnych algorytmów dla stosunkowo prostych procesów jednomodowych, jak również bardziej złożonych systemów wielomodowych. Zwartość rozprawy, opisana w poprzednim punkcie recenzji, wynika bezpośrednio z postawionego celu i koncentruje się na opracowaniu, a następnie analizie zbudowanego narzędzia.

Do badania procesów jednomodowych Autorka zaproponowała nowy, interesujący algorytm detekcji anomalii z klasyfikatorem binarnym (nazwany BaFFle). Metoda, oparta na analizie PCA i dwóch alternatywnych metodach estymacji rozkładu gęstości prawdopodobieństwa, wydaje się być

R. Jan

dość prosta. Jednak interesującym pomysłem jest wprowadzenie w niej mechanizmu adaptacyjnej zmiany limitów poziomu sygnałów w trakcie pracy systemu. Wyniki przedstawione w rozdziale praktycznym pokazują, że algorytm adaptacyjnej zmiany limitu sygnałów istotnie poprawia skuteczność detekcji anomalii. Metoda ta jest szczególnie efektywna w połączeniu z estymacją gęstości estymatorem jądrowym (ang. *Kernel Density Estimation* – KDE).

Do wykrywania anomalii procesów wielomodowych Autorka zastosowała algorytm klasteryzacji oparty na procesie Dirichleta z modelem mieszanin gaussowskich (DP-GMM). Rezultat w postaci finalnego algorytmu grupowania, zaimplementowanego w systemie wykrywania anomalii, oceniam jak najbardziej pozytywnie. Jednak należy podkreślić, że wspomnianą implementację poprzedziła dogłębna analiza teoretyczna i symulacyjna zastosowanej metody. Zdaniem recenzenta, szczegółowa i krytyczna analiza metody DP-GMM stanowiąca podrozdziały 4.4 do 4.6 jest istotną kontrybucją Doktorantki.

Najbardziej zaawansowany system detekcji anomalii w procesach wielomodowych został przedstawiony w rozdziale piątym. Opracowana przez Doktorantkę autorska koncepcja opiera się na modelu stanów procesu w oparciu o autoregresyjny model w postaci równań stanu oraz estymatorze będącym pewnym uogólnieniem filtru Kalmana (nazwanym FKF). Autorka w ramach pracy zbudowała podstawy teoretyczne systemu, przygotowała algorytm działania oraz dokonała jego implementacji. Zarówno badania symulacyjne, jak również badania praktyczne na podstawie danych z wielofazowej instalacji przepływowej, wykazały wysoką skuteczność algorytmu. Zadaniem recenzenta to oryginalne rozwiązanie stanowi najistotniejsze osiągnięcie recenzowanej rozprawy. O dużej wadze przedmiotowego osiągnięcia świadczą artykuły Doktorantki, które zostały opublikowane w prestiżowych czasopismach naukowych.

Finalnie, przedstawione w pracy doktorskiej rozwiązania doprowadziły do opracowania systemu pozwalającego na monitorowanie procesów o różnym stopniu złożoności. Tym samym cel pracy został osiągnięty.

Podsumowując ocenę merytoryczną należy zaznaczyć, że główny wkład Kandydatki w rozwój dyscypliny Automatyka, Elektronika i Elektrotechnika został przedstawiony w rozdziałach trzecim, czwartym i piątym. Przy czym szczególnie istotny, zdaniem recenzenta, są wyniki przedstawione w rozdziale piątym. Również ważny w ujęciu praktycznym jest rozdział szósty, który weryfikuje opracowane metody w oparciu o rzeczywiste dane.

3. Analiza źródeł, pozycja rozprawy, znaczenie wyników Autora, umiejętność przedstawiania wyników

Motywacja dla podjęcia tematu rozprawy wynika z dobrze przeprowadzonej przez Autorkę analizy literatury przedmiotu, liczącej 238 pozycji. Dzięki szerokiej analizie został poprawnie odzwierciedlony aktualny stan wiedzy na temat zagadnień związanych z wykorzystywanymi w pracy metodami wykrywania uszkodzeń w oparciu o wnioskowanie statystyczne.

Należy zauważyć, że pozycja rozprawy w stosunku do stanu wiedzy reprezentowanej w literaturze światowej jest wyróżniająca. Warte szczególnego podkreślenia jest tu współautorstwo Doktorantki w sześciu cennych artykułach z zakresu recenzowanej dysertacji, z czego trzech w renomowanych czasopismach tj, (*EURASIP Journal on Advances in Signal Processing*, *Journal of Process Control*, *IEEE Transactions in Control System Technology*).



Autorka posiadała umiejętność poprawnego, przekonywującego i precyzyjnego przedstawiania uzyskanych przez siebie wyników. Zarówno rozprawę, jak również artykuły cechuje zwięzłość, jasność języka oraz precyzja. Należy również podkreślić systematyczność pracy Doktorantki w poszczególnych rozdziałach dysertacji. Autorka każdy rozdział merytoryczny pracy zaczyna od elementów podstawowych, a następnie stopniowo przeprowadza rozważania nad kolejnymi elementami opracowywanych metod. Zdecydowanie ułatwia to czytanie pracy i interpretację wyników.

4. Główne wady rozprawy, słabe strony, uwagi i pytania

Należy podkreślić, że recenzowana rozprawa doktorska napisana jest bardzo precyzyjnie oraz przedstawia szereg nowych, interesujących i ważnych w ujęciu naukowym zagadnień. Poziom merytoryczny rozprawy wymagał od recenzenta dużego zaangażowania w czytaniu i weryfikacji prezentowanych wyników. Jednakże, pomimo braku istotnych wad i niedostatków, warto wskazać na pewne, w większości drobne uchybienia oraz niejasności. Ponadto, w trakcie poznawania pracy recenzentowi nasunęły się pytania ogólne. Zostały one podzielone na dwie grupy, a) uwagi dyskusyjne i pytania oraz b) drobne uchybienia typograficzne i stylistyczne.

Uwagi dyskusyjne i pytania:

- W dysertacji występuje bardzo dużo skrótów (łącznie jest ich kilkadziesiąt). Autorka każdy skrót poprzedza pełną nazwą przy pierwszym użyciu, jednak tak duża ich liczba zdecydowanie utrudnia czytanie pracy. Wprowadzenie syntetycznego opisu skrótów na początku pracy zdecydowanie ułatwiłoby jej czytanie.
- Niektóre fragmenty, szczególnie w rozdziale drugim mają bardzo ogólny charakter. Poza wysokim stopniem ogólności który, zdaniem recenzenta, nie jest potrzebny, rozdział zawiera elementy dość luźno związane z pozostałą częścią pracy, np. podrozdział 2.2 Methodologies of process modelling.
- Algorytm prezentujący działanie algorytmu BaFFle w rozdziale trzecim (Rys. 3.3) nie pokazuje procesu estymacji rozkładu gęstości sygnału. Zdaniem recenzenta ten proces jest dość istotny w całym algorytmie i powinien być umieszczony.
- Adaptacja zmian parametrów w metodzie BaFFle (równanie (3.23) i (3.24)) odbywa się metodą najmniejszych średnich kwadratów (ang. *Least Mean Squares*). Czy implementacja w tym miejscu adaptacyjnej metody najmniejszych kwadratów (ang. *Adaptive Least Squares*) poprawiłaby działanie algorytmu?
- Rozpatrywana w rozdziale piątym metoda opiera się na modelu liniowym. Jednak w praktyce przemysłowej zwykle spotka się procesy nieliniowe. W jaki sposób Doktorantka rozważa rozszerzenie proponowanej metody na procesy nieliniowe?

Drobne uchybienia typograficzne i stylistyczne:

- W spisie treści dysertacji nie zawarto bibliografii.

Roh

- W Tabeli 6.3 zabrakło określenia jednostek przedstawianych wyników. Do zrozumienia wyników przedstawionych w tabeli niezbędne jest przeczytanie komentarza Autorki.
- Strona 24, linia 11-12; Układ graficzny strony sprawia wrażenie, że czegoś brakuje między linią 11 a 12.
- Strona 54, linia 12; Brak spacji między wyrazami.

Należy podkreślić, że przedstawione powyżej uchybienia i uwagi, mają w większości charakter dyskusyjny i nie obniżają w żaden sposób bardzo pozytywnej oceny pracy.

5. Podsumowanie recenzji i Wniosek końcowy

Recenzowana rozprawa doktorska stanowi, zdaniem recenzenta, oryginalne rozwiązanie ważnego problemu naukowego oraz wykazuje dużą ogólną wiedzę teoretyczną i aplikacyjną Kandydatki w dyscyplinie naukowej Automatyka, Elektronika i Elektrotechnika, a także Jej umiejętność samodzielnego prowadzenia pracy naukowej. Zatem na podstawie art. 179 Ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1669) stwierdzam, że **rozprawa ta spełnia z nadmiarem** warunki określone w art. 13 ust. 1 i ust. 2 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. 2003 poz. 595, z późniejszymi zmianami).

Uwzględniając oryginalność rozwiązania problemu naukowego przedstawionego w rozprawie, specjalistyczną wiedzę Kandydatki w dyscyplinie Automatyka, Elektronika i Elektrotechnika oraz umiejętność samodzielnego prowadzenia pracy naukowej **wnoszę o dopuszczenie rozprawy doktorskiej mgr inż. Tian Cong do publicznej obrony.**

Rafał Staniśkowski