

Warszawa, 20.10.2020

Prof. dr hab. inż. Jan Maciej Kościelny
Instytut Automatyki i Robotyki
Wydział Mechatroniki
Politechnika Warszawska
ul. św. A. Boboli 8
02-525 Warszawa
tel. kom. 601 810 419
e-mail: jmk@mchtr.pw.edu.pl

**Recenzja pracy doktorskiej
mgr inż. Artura Maździarza**

Wykorzystanie ciągów binarnych w metodach korelacji alarmów w mobilnych sieciach telekomunikacyjnych

1. Tematyka rozprawy

Tematem rozprawy są zagadnienia diagnostyki mobilnych sieci telekomunikacyjnych. W pracy rozwijane są metody korelacji alarmów. Celem działań diagnostycznych jest generowanie hipotez przyczyn awarii na podstawie automatycznie wykrywanych wzorców sekwencji alarmów. Współczesne sieci telefonii komórkowej, charakteryzują się dużą złożonością. Duże zbiory danych, które muszą być analizowane w czasie rzeczywistym podczas diagnostyki awarii mobilnych sieci telekomunikacyjnych wymagają automatyzacji zadań korelacji alarmów i analizy przyczyn awarii.

Diagnostyka systemów technicznych stanowi szybko rozwijającą się gałąź wiedzy inżynierskiej. Rozwijane są zarówno elementy teorii jak i podejścia praktyczne, zorientowane na zastosowania w rzeczywistych instalacjach technologicznych. Współczesne systemy diagnostyczne bazują na sformalizowanych teoriach wywodzących się z automatyki, analizy systemowej czy inżynierii wiedzy (sztucznej inteligencji). Mogą one także opierać się na wykorzystaniu wiedzy eksperckiej i doświadczenia.

Czynnikami warunkującymi konieczność rozwoju metod monitorowania i diagnostyki są przede wszystkim wymagania dotyczące bezpieczeństwa i niezawodności coraz bardziej złożonych systemów oraz wymagania ekonomiczne dotyczące unikania strat związanych z awariami. Wczesna i precyzyjna diagnoza umożliwia szybkie usuwanie zaistniałych uszkodzeń i zapobieganie wystąpieniu groźnych awarii, a w konsekwencji pozwala na redukcję strat ekonomicznych związanych z uszkodzeniami. Automatycznie realizowana diagnostyka jest także elementem systemów sterowania umożliwiającym tolerowanie uszkodzeń. Rozwój metod i systemów diagnostycznych ma istotne znaczenie dla efektywnego i bezpiecznego użytkowania złożonych instalacji technologicznych i jest jednym z istotnych czynników warunkujących konkurencyjność i szeroko rozumianą niezawodność systemów.

Tematyka recenzowanej rozprawy wpisuje się zatem dobrze w aktualne trendy badań naukowych. Duża ilość alarmów w sieci telekomunikacyjnej wymaga stosowania algorytmów

redukujących nadmiarowe informacje i wyznaczających związki przyczynowo-skutkowe między zdarzeniami w celu szybkiej identyfikacji przyczyn awarii i neutralizacji problemów w sieci.

Uszkodzenia w sieciach telekomunikacyjnych generują dużą ilość alarmów. W przypadku awarii sieć komórkowa może wygenerować kilkanaście, a nawet kilkadziesiąt alarmów na sekundę. Prawidłowa eksploatacja sieci wymusza konieczność szybkiego diagnozowania awarii i ich usuwania. Natomiast obecnie brak jest automatycznych metod korelacji alarmów, które spełniałyby wszystkie wymagania oraz były proste do zastosowania. Celowe są zatem badania ukierunkowane na opracowanie efektywnych metod korelacji alarmów oraz generowanie hipotez przyczyn awarii dla dużych zbiorów danych.

2. Cel i zakres pracy

Celem pracy było opracowanie szybkich i skutecznych metod korelacji alarmów, które będą użyteczne w zarządzaniu sieciami telekomunikacyjnymi przy dużych zbiorach alarmów. Założono, że metoda powinna umożliwiać określanie kierunku relacji przyczynowo-skutkowej oraz uwzględniać opóźnienia między symptomami a skutkiem.

Sformułowano następującą tezę pracy: *wykorzystanie ciągów binarnych do reprezentacji atrybutów alarmów pozwala na tworzenie algorytmów korelacji dużych zbiorów alarmów, o krótkim czasie wykonania obliczeń w zadanym oknie czasowym korelacji oraz możliwością wyznaczenia siły i kierunku związku między alarmami z uwzględnieniem efektu propagacji (opóźnienia) między alarmem źródłowym a powiązany z nim efektem, skutkiem.*

3. Zawartość i struktura rozprawy

Praca liczy 145 stron i zawiera: spis treści, streszczenie w języku polskim i angielskim, wykaz oznaczeń, siedem rozdziałów, wykaz skrótów oraz wykaz literatury.

Rozdział pierwszy jest wstępem do rozprawy i zawiera cel oraz tezę pracy. W rozdziale drugim przedstawiono przegląd zagadnień związanych z zarządzaniem mobilnymi sieciami telekomunikacyjnymi ze szczególnym uwzględnieniem metod korelacji alarmów. Podano charakterystykę mobilnych sieci telekomunikacyjnych, scharakteryzowano topologie sieci, wprowadzono pojęcia z zakresu korelacji zdarzeń oraz scharakteryzowano stan badań w zakresie tych technik diagnostyki defektów.

W rozdziale trzecim zaprezentowano rozwijane w pracy metody korelacji alarmów wykorzystujące ciągi binarne, czyli metody korelacji czasów wystąpienia alarmów oraz metody analizy skupień. Opracowana metoda korelacji alarmów umożliwia wykrycie związku przyczynowo-skutkowego między alarmami z dynamicznym oszacowaniem rozmiaru okna korelacji alarmów. Zaproponowana metoda analizy skupień realizowana jest techniką k-średnich z filtrowaniem topologicznym.

Wyniki badań eksperymentalnych prowadzonych dla rzeczywistej sieci telekomunikacyjnej zestawiono w rozdziale czwartym. W badaniach wykorzystano dane alarmowe z sieci zawierającej elementy sieci technologii 2G, 3G i 4G. Dane użyte w symulacjach zawierały informacje o 1440813 zdarzeniach alarmowych zebranych w okresie od 07.2014 r. do 05.2015 r. Dla każdego alarmu zarejestrowanych zostało pięć atrybutów: czas wystąpienia,

numeryczny identyfikator i priorytet alarmu, opis alarmu oraz identyfikator elementu, który wygenerował alarm. Do symulacji użyte zostało środowisko pakietu R w wersji 3.3.1.

Rozdział piąty zawiera dyskusję uzyskanych wyników wraz z analizą krytyczno-porównawczą opracowanych metod w zestawieniu z innymi technikami stosowanymi do korelacji alarmów. Wnioski oraz cele dalszych prac są przedstawione w końcowych rozdziałach szóstym i siódmym.

Wykaz cytowanej literatury zawiera 119 pozycji, w tym najważniejsze publikacje związane z tematem rozprawy. Można stwierdzić, że jest to solidnie opracowany i aktualny przegląd literatury dotyczącej metod korelacji alarmów. Świadczy on o dobrym rozpoznaniu stanu badań w zakresie tematyki pracy doktorskiej. Ponadto 5 cytowanych prac, to publikacje, których autorem jest doktorant.

4. Oryginalne wyniki rozprawy

Osiągnięcia Autora rozprawy oraz jej oryginalne elementy dotyczą opracowania dwóch metod korelacji alarmów przeznaczonych dla mobilnych sieci telekomunikacyjnych. Przy wyborze metod korelacji alarmów brano pod uwagę wiedzę ekspercką wskazującą na kluczowe znaczenie czasu występowania alarmów w sieciach telekomunikacyjnych. Zaobserwowano, że skorelowane alarmy w mobilnej sieci telekomunikacyjnej mogą pojawiać się w tym samym czasie lub w przedziale od jednej do czterech sekund. Dlatego zaproponowane metody korelacji alarmów odnoszące się do czasu ich występowania.

Pierwsza z opracowanych metod korelacji alarmów polega na analizie ciągów binarnych. Metoda pozwala na wykrycie związku przyczynowo-skutkowego między alarmami. Do określenia tego związku wykorzystywana jest analiza współczynników podobieństwa ciągów binarnych Dice'a, Dice1 oraz Dice2 dla binarnej reprezentacji czasów wystąpienia alarmów. Ponadto w sposób dynamiczny szacowany jest rozmiar okna korelacji alarmów z zastosowaniem odległości Hamminga. Istotnym problemem, który wymagał rozwiązania było podanie sposobu ustalenia rozmiaru okna czasowego zależnego od rodzaju elementów sieci telekomunikacyjnej i jej architektury. Dokładność określenia rozmiaru okna czasowego wpływa na dokładność procesu korelacji alarmów, jak również na wydajność implementowanych metod korelacji. Metoda pozwala na określenie siły i kierunku związku między analizowanymi alarmami oraz uwzględnia opóźnienie między przyczyną wystąpienia zdarzenia a skutkiem.

Druga opracowana metoda korelacji alarmów wykorzystuje analizę skupień (klasyfikację bez nadzoru) przy użyciu metody k-średnich. Celem jest wykrywanie sekwencji alarmów, które wystąpiły po sobie w krótkim odstępie czasu w zbiorze danych alarmowych. Alarmy przyporządkowywane są do K skupień, przy czym dana obserwacja jest przypisana do skupienia (zbioru) o najbliższej jej wartości średniej. W celu redukcji danych stosowane jest filtrowanie topologiczne, polegające na eliminacji danych pochodzących z innych niż analizowany łańcuchów topologii sieci. Podział alarmów na klastry umożliwia formułowanie hipotez o przyczynach sekwencji alarmów w skupieniu. Przyczyna związana jest zwykle z pierwszym alarmem w klastrze. Hipotezy przyczyny awarii weryfikowane są przez ekspertów.

Podkreślić należy, że opracowane metody były badane eksperymentalnie. Wykorzystano dane zebrane z heterogenicznej mobilnej sieci telekomunikacyjnej zawierającej elementy technologii 2G, 3G i 4G z okresu 07.2014 – 05.2015. Wybrane próbki danych do analizy

zawierały alarmy pochodzące z sieci zawierającej 28 elementów typu BSC (kontroler stacji bazowej 2G) i 27 typu RNC (kontroler sieci radiowej 3G). Podano dwa przykłady dla metody „ruchomy Dice” oraz jeden przykład dla metody analizy skupień. W przypadku tej drugiej metody w celu minimalizacji czasów przetwarzania stosowano wstępne przetwarzanie polegające na grupowaniu alarmów w zbiory pochodzące z jednego łańcucha topologii. Przeprowadzono także wartościową dyskusję krytyczno-porównawczą opracowanych metod z innymi metodami korelacji alarmów.

Przeprowadzone badania potwierdziły, że możliwe jest opracowanie algorytmów korelacji dużych zbiorów alarmów, o krótkim czasie wykonania obliczeń w zadanym oknie czasowym korelacji oraz istnieje możliwość wyznaczenia siły i kierunku związku między alarmami z uwzględnieniem efektu propagacji (opóźnienia) między alarmem źródłowym a powiązanim z nim efektem, skutkiem. Teza pracy została zatem udowodniona.

5. Uwagi krytyczne i dyskusyjne.

Poniższe uwagi krytyczno-dyskusyjne są formułowane z pozycji specjalisty w zakresie diagnostyki procesów przemysłowych i układów sterowania automatycznego, czyli osoby nie związanej z technikami sieci telekomunikacyjnych. Jednak pewne obszary dotyczące analizy alarmów w systemach monitorowania procesów SCADA i systemów sterowania klasy DCS są podobne do zagadnień korelacji alarmów w sieciach telekomunikacyjnych. Należy do nich problem nadmiaru alarmów generowanych w sytuacjach wystąpienia groźnych uszkodzeń. Podobnie jak w sieciach telekomunikacyjnych w systemach sterowania procesami powstaje bardzo dużo alarmów (nawet ponad 500 w okresie 1-2 minut). Przy takiej liczbie alarmów pojawia się problem przeciążenia informacyjnego operatorów, którzy działają w warunkach stresu i nie są w stanie podjąć właściwych decyzji zabezpieczających nadzorowany obiekt.

Nadmiar alarmów jest tylko jedną z wad systemów alarmowych. Inną poważną wadą, która cechuje zarówno aplikacje tych systemów w obszarze układów sterowania jak też sieci telekomunikacyjnych jest brak realizacji w sposób automatyczny fazy zaawansowanej lokalizacji uszkodzeń (fault isolation), której wynikiem są diagnozy wskazujące uszkodzenia.

Redukcja nadmiaru alarmów jest zatem działaniem racjonalnym, ale pojawia się pytanie czy jest to najlepsze podejście do budowy systemów automatycznego nadzoru sieci telekomunikacyjnych?

W przypadku diagnostyki zautomatyzowanych obiektów przemysłowych od ponad trzydziestu lat rozwijane są metody detekcji uszkodzeń wykorzystujące modele cząstkowe diagnozowanego obiektu. Modele takie reprezentują stan normalny, bez uszkodzeń, a symptomy uszkodzeń są pozyskiwane na drodze generacji residuów, czyli rozbieżności między sygnałami mierzonymi i wzorcowymi, którymi są wyjścia modeli. Takie podejście do detekcji w połączeniu z automatycznym wnioskowaniem pozwala uzyskiwać szybkie i precyzyjne diagnozy nawet w przypadku uszkodzeń o małych rozmiarach. Diagnozy znacznie dokładniej określają zaistniałe uszkodzenia niż sekwencje alarmów. Pojawia się zatem pytanie, czy analogicznego podejścia nie można zastosować także w przypadku sieci telekomunikacyjnych? Liczę na ustosunkowanie się Doktoranta do tego pytania w trakcie obrony pracy.

W rozdziale 5, dotyczącym dyskusji związanej z porównaniem opracowanej metody z innymi, Autor stwierdza, że brak jest formalnie zdefiniowanych kryteriów takiego porównania. Stwierdzenie to jest kontrowersyjne. Wynika z rozpatrywania jedynie metod analizy alarmów. Jednak metody analizy alarmów powinny być traktowane jako metody diagnostyki. Celem diagnostyki jest szybkie i niezawodne wykrywanie i rozróżnianie uszkodzeń (fault detection and isolation). Zatem porównanie metod powinno być przeprowadzone w aspekcie wskaźników wykrywalności i rozróżnialności uszkodzeń. Natomiast np. poziom redukcji alarmów czy dokładność metody odnosząca się do poprawności wyznaczenia kierunku relacji nie mają charakteru ogólnego, lecz odnoszą się jedynie do metod korelacji alarmów.

6. Podsumowanie

Recenzowana praca doktorska lokuje się w dziedzinie nauk inżyniersko-technicznych w dyscyplinie informatyka techniczna i telekomunikacja. Na podstawie lektury pracy można stwierdzić, że Autor jest specjalistą w zakresie metod korelacji alarmów w sieciach telekomunikacyjnych oraz dociekliwym naukowcem, umiejącym rozwiązywać sformułowane problemy badawcze. O wysokich kompetencjach Autora świadczą zarówno oryginalne elementy pracy jak też dobra orientacja w stanie badań w zakresie tematyki rozprawy, sposób rozumowania, umiejętność formułowania wniosków, solidność prowadzonych analiz.

Podniesione wyżej uwagi dyskusyjne nie wpływają na ogólną pozytywną ocenę pracy, zawierającej oryginalne osiągnięcia naukowe oraz wartościowy materiał badawczy. Potencjał użyteczności praktycznej uzyskanych wyników jest znaczący. Autor wykazał umiejętność rozwiązywania trudnych problemów badawczych i technicznych.

W związku z powyższym stwierdzam, że recenzowana rozprawa doktorska pana mgr. inż. Artura Maździarza spełnia wszystkie wymagania stosownej ustawy o tytułach i stopniach naukowych. Wnoszę zatem o przyjęcie opiniowanej rozprawy oraz dopuszczenie jej do publicznej obrony przez Radę Naukową Instytutu Badań Systemowych PAN.

