

Warszawa, 9 stycznia 2025

prof. dr hab. Anna Gambin
Instytut Informatyki
Uniwersytetu Warszawskiego

RECENZJA ROZPRAWY DOKTORSKIEJ

mgra inż. Bartłomieja Szawulaka

ZATYTUŁOWANEJ

METODY PORÓWNYWANIA MODELI ZŁOŻONYCH SYSTEMÓW BIOLOGICZNYCH WYRAŻONYCH ZA POMOCĄ SIECI PETRIEGO

1. Problem badawczy i jego znaczenie

Praca doktorska dotyczy problemu porównywania sieci Petriego, które są kluczowym narzędziem w modelowaniu procesów biologicznych i innych systemów dynamicznych. W teorii sieci Petriego istnieje pojęcie bisymulacji, które pozwala na porównywanie zachowań systemów modelowanych za pomocą tych sieci. Bisymulacja umożliwia wykazanie równoważności dwóch systemów.

Zadanie jakiego podejmuje się Autor rozprawy, to konstrukcja miar podobieństwa dla sieci Petriego, które nie opisują równoważnych systemów, a jedynie systemy o zbliżonej semantyce. Trzeba podkreślić, że porównywanie sieci Petriego jako struktur grafowych sprowadza się do analizy składniowej, czyli badania ich formy z pominięciem znaczenia funkcjonalnego, co może potencjalnie ograniczać praktyczną przydatność tych metod w kontekście biologii systemów i innych dziedzin.

Zasadniczym problemem badawczym jest zatem opracowanie metod porównywania, które uwzględniają zarówno składnię, jak i semantykę sieci Petriego, co umożliwi ich zastosowanie w biologii systemów, gdzie sieci służą do modelowania procesów, takich jak ścieżki metaboliczne, regulacja genów czy sygnalizacja komórkowa. Jest oczywiste, że skuteczne metody porównywania sieci mogą znacząco usprawnić badania nad złożonością systemów opisywanych przez nie.

Praca wnosi istotny wkład w dziedzinę, proponując nowe metody porównywania sieci Petriego, które wykraczają poza standardowe podejścia grafowe. Wprowadzone rozwiązania, inspirowane istniejącymi technikami porównywania grafów, zostały dostosowane do specyfiki sieci Petriego. Kluczowym punktem jest rozwój metod umożliwiających dekompozycję sieci na podsieci o biologicznym znaczeniu, co pozwala na bardziej precyzyjne ujęcie

semantycznych różnic między modelami.

Poruszone w rozprawie zagadnienia mają nie tylko znaczenie teoretyczne, lecz również praktyczne, wspierając rozwój narzędzi obliczeniowych w biologii systemów. Warto zaznaczyć, że istniejące w literaturze uogólnienia wspomnianego pojęcia bisymulacji do postaci miary podobieństwa nie znalazły do tej pory zastosowań.

2. Wkład Autora

Główne dokonania Autora w ramach rozprawy doktorskiej koncentrują się na projektowaniu, implementacji oraz testowaniu metod porównywania sieci Petriego w kontekście modeli biologicznych. Praca obejmuje rozwój nowych metod analitycznych oraz adaptację istniejących technik porównywania grafów do specyficznej struktury sieci Petriego. Osiągnięcia autora znajdują wyraz zarówno w publikacjach naukowych, jak i w rozwoju narzędzi wspierających analizy systemów biologicznych.

Autor opracował metody porównywania sieci Petriego bazujące na grafletach, w tym nowe struktury pn-grafletów. Przeprowadził szczegółowe analizy wrażliwości rozkładu stopni grafletów na gęstość sieci, co pozwoliło na identyfikację sytuacji, w których metody te wykazują najlepszą skuteczność. Dodatkowo Autor wprowadził metodę porównywania opartą na dekompozycji sieci do podsieci o znaczeniu biologicznym, co umożliwia precyzyjną ocenę podobieństwa modeli z punktu widzenia ich funkcji biologicznych.

W rozprawie przedstawiono algorytmy bazujące na definicji wierzchołków rozgałęziających, które stanowią kluczowy element metody dekompozycji. Autor przedstawił również wyniki szerokich testów metod porównywania, wskazując ich zalety i ograniczenia oraz proponując kierunki dalszych usprawnień.

Bartłomiej Szawulak jest autorem lub współautorem kilku istotnych publikacji związanych z rozprawą. Do najważniejszych należą artykuły w *Scientific Reports* („Graphlets in comparison of Petri net-based models of biological systems”), a także prace w recenzji dotyczące porównywania sieci („Comparison by partition - finding Petri nets similarity on the basis of subnets”) oraz wrażliwości metryk grafletowych („Graphlet degree distribution sensibility in Petri net with low density”).

Ponadto, Autor brał udział w projektach takich jak tworzenie narzędzia Holmes (publikacja w *Bioinformatics* pt. „Holmes: a graphical tool for development, simulation and analysis of Petri net based models of complex biological systems”), umożliwiającego graficzne projektowanie, symulację i analizę modeli opartych na sieciach Petriego, co dodatkowo podkreśla jego wkład w rozwój narzędzi wspierających badania nad złożonymi systemami biologicznymi.

3. Poprawność

Autor przedstawia jasne uzasadnienia i przekonujące motywacje dla wprowadzanych metod i wykorzystuje różnorodne podejścia do analizy porównawczej sieci Petriego, w tym dekompozycję do podsieci, graflety oraz nowe koncepcje, takie jak wierzchołki rozgałęziające. Eksperymenty obliczeniowe zostały zaplanowane z uwzględnieniem specyfiki biologicznych zastosowań, a wybór adekwatnych przykładów, opisanych szczegółowo w Rozdziałach 6-10, podkreśla praktyczne znaczenie metod. Przykłady te obejmują modele o zróżnicowanej strukturze, co zwiększa wiarygodność wyników oraz daje nadzieję na ich generalizację.

Jednym z aspektów pracy, które mogłyby zostać rozwinięte, jest analiza statystyczna otrzymanych wyników porównań. Kluczowe jest pojęcie p-wartości, które może posłużyć do oceny istotności statystycznej. Uwzględnienie tej miary pozwoliłoby bardziej obiektywnie interpretować rezultaty eksperymentów i osadzić je w solidnym kontekście statystycznym. Wprowadzenie p-wartości mogłoby zostać oparte na modelu permutacyjnym, w którym porównywane sieci byłyby losowo przekształcane w celu oszacowania rozkładu wyników i wyznaczenia istotności uzyskanych różnic. Alternatywnie, można rozważyć analityczne podejście, które umożliwiłoby bezpośrednie wyprowadzenie p-wartości w oparciu o założenia teoretyczne dotyczące rozkładu miar porównawczych.

Warto również zauważyć, że dla pewnych wskaźników, takich jak indeks Tanimoto, istnieją w literaturze metody obliczania p-wartości, co można wykorzystać w kontekście porównywania sieci Petriego z użyciem tego indeksu. Włączenie p-wartości nie tylko podniosłoby wiarygodność uzyskanych wyników, ale także ułatwiłoby ich interpretację w szerszym kontekście naukowym, szczególnie w badaniach, które wymagają rygorystycznej analizy statystycznej.

Inną potencjalną możliwością jest zbadanie przydatności metryk opartych na pojęciu bisymulacji w kontekście modeli biologii systemów. Taka miara, podobnie jak wykorzystanie niezmienników, odnosi się do semantyki modelowanego układu, a nie jedynie do struktury sieci, która może być porównana do składni kodu w językach programowania. Dwa programy, nawet jeśli składają się z takiej samej liczby konstrukcji if-then-else oraz analogicznie zbudowanych zagnieżdżonych pętli, mogą rozwiązywać odmienne problemy.

Naturalnym rozszerzeniem stosowanym w biologii systemów są stochastyczne sieci Petriego, gdzie niedeterminizm zostaje zastąpiony przez prawdopodobieństwo odpalenia tranzycji. Stochastyczne sieci Petriego są równoważne dyskretnym procesom Markowa. Ciekawa byłaby dyskusja nad technikami porównywania takich sieci.

Podsumowując, wprowadzone przez Autora podejścia do porównywania sieci Petriego stanowią ważny, choć wstępny krok w tej tematyce. Doceniam zaprezentowane w pracy

porównanie nowych rozwiązań z istniejącymi podejściami, jak izomorfizm grafów czy wykorzystanie niezmienników, co podnosi spójność merytoryczną rozprawy. Warto podkreślić, że wyniki eksperymentalne są zaprezentowane w sposób klarowny i poparte dyskusją, która uwzględnia zarówno mocne strony, jak i ograniczenia proponowanych metod.

4. Wiedza kandydata

Rozdziały 2, 3, i częściowo 4 rozprawy w największym stopniu prezentują istniejący stan wiedzy, co pozwala ocenić ogólną wiedzę kandydata w zakresie informatyki, w szczególności w obszarze teorii grafów, systemów formalnych oraz metod porównywania sieci i grafów.

Rozdział 2 wprowadza podstawowe pojęcia z teorii grafów oraz definicje klasycznego wariantu sieci Petriego, w tym jej rozwój historyczny i znaczenie w modelowaniu systemów. Jest to kompleksowy przegląd podstawowych zagadnień teoretycznych, który dobrze podkreśla zrozumienie przez kandydata podstawowych elementów tej dziedziny. Rozdział ten wyczerpująco wprowadza czytelnika w kluczowe koncepcje, co świadczy o solidnym przygotowaniu merytorycznym.

Rozdział 3 przedstawia szczegółowy przegląd metod porównywania grafów, w tym takich technik jak izomorfizm grafów, dystans grafowy, największy wspólny podgraf czy grafowy dystans edycyjny (GED). Opisane są także zaawansowane podejścia, takie jak graflety i ich różne metryki podobieństwa (RGF, GDD, GDDA). Te sekcje wskazują, że kandydat dobrze zna metody analizy grafów stosowane w praktyce i potrafi ocenić ich zastosowanie w różnych kontekstach.

Rozdział 4 omawia istniejące metody dekompozycji sieci Petriego, co pokazuje znajomość technik stosowanych do analizy systemów formalnych.

Bibliografia rozprawy jest obszerna i adekwatna, zawierając odniesienia do głównych prac w dziedzinie grafów, sieci Petriego oraz metod ich analizy, a także do współczesnych badań w tej tematyce.

Jakość rozprawy oceniam wysoko. Jest napisana klarownym językiem, cechuje się logiczną strukturą oraz zwięzłym przedstawieniem wprowadzanych koncepcji. Jej treść wskazuje, że Autor posiada ogólną wiedzę w zakresie informatyki technicznej i telekomunikacji, zwłaszcza w kontekście teorii grafów i systemów modelowania. Doceniam prezentowanie algorytmów w postaci pseudokodu. Jedynie w niektórych przypadkach brakowało mi podsumowania złożoności obliczeniowej i pamięciowej omawianych rozwiązań.

5. Inne uwagi

Poniższe drobne uwagi, pytania i sugestie poprawek nie zmieniają mojego pozytywnego zdania o rozprawie.

1. Rozdział 2 w opisie historycznym sieci Petriego zawiera nieścisłość: dopiero rozszerzenie sieci o inhibitory (odpowiadające tzw testowi zero) pozwala na symulowanie Maszyny Turinga.
2. Wzór (3.19) definiujący podobieństwo enzymów jest niezrozumiały.
3. Definicja sieci funkcyjnej w Rozdziale 4.5 nie jest wystarczająco jasna.
4. Na jakiej podstawie wyznaczono parametry typowych sieci do testowania w Rozdziale 6.5.2 ?
5. Użycie koloru na Rysunkach 6.19 i 6.20 zwiększyłoby czytelność wykresów.
6. Jakie zjawiska biologiczne modelują sieci badane w Tabeli 6.8 ?

8. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez artykuł 13 Ustawy z dnia 14 marca 2003 r. o stopniach i tytule naukowym (z późniejszymi zmianami¹), stwierdzam, że recenzowana przeze mnie praca spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy. Praca zawiera oryginalne rozwiązanie problemu naukowego, co w pełni wpisuje się w wymogi formalne. Jednocześnie, analiza jej treści pozwala jednoznacznie stwierdzić, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka techniczna i telekomunikacja oraz wykazuje umiejętność samodzielnego prowadzenia pracy naukowej, co jest podstawą oceny jego kompetencji badawczych. W związku z tym wnoszę o dopuszczenie magistra Bartłomieja Szawulaka do dalszych etapów postępowania w sprawie nadania stopnia doktora.

¹<https://isap.sejm.gov.pl/isap.nsf/home.xsp>