



Akademia Górniczo-Hutnicza  
im. Stanisława Staszica w Krakowie  
Wydział Fizyki i Informatyki Stosowanej  
al. Mickiewicza 30, 30-059 Kraków

dr hab. inż. Małgorzata Krawczyk  
tel. +48 12 617 35 18  
malgorzata.krawczyk@agh.edu.pl

Kraków, 14 lipca 2016

**Recenzja rozprawy doktorskiej mgr. inż. Maksymiliana Bujoka  
pt. „Teoria, symulacje numeryczne i wybrane zastosowania  
sieci złożonych o strukturze blokowej”.**

Praca doktorska mgr. inż. Maksymiliana Bujoka pt. „Teoria, symulacje numeryczne i wybrane zastosowania sieci złożonych o strukturze blokowej” wykonana została na Wydziale Fizyki Politechniki Warszawskiej pod kierunkiem dr hab. Agaty Fronczak. Podstawę prezentowanej rozprawy stanowią trzy publikacje, których Doktorant jest współautorem. Jedna z nich ukazała się w *Physical Review E* w roku 2013, druga w *European Physical Journal B* w roku 2014 i ostatnia, także w roku 2014, w *Acta Physica Polonica B Proceedings Supplement*. Zawarte w nich wyniki prezentowane były na dwóch konferencjach międzynarodowych w formie czterech wystąpień bądź plakatów, w tym dwóch przez mgr. Bujoka. Manuskrypt rozprawy obejmuje pięć rozdziałów, z czego pierwszy zawiera sformułowanie problemu badawczego a ostatni podsumowanie oraz nakreślenie kierunku ewentualnych dalszych badań. Zasadniczą część stanowią trzy rozdziały dotyczące odpowiednio modelu sieci blokowych, perkolacji na sieciach blokowych oraz analizy sieci semantycznej dla języka polskiego i angielskiego.

Praca podejmuje rozwijaną szeroko w ostatnich latach tematykę badania układów złożonych, które bardzo często są opisywane i modelowane za pomocą grafów. Opis taki z jednej strony wyraża istnienie powiązań pomiędzy określonymi elementami wchodzącymi w skład analizowanych układów, z drugiej - umożliwia badanie ich własności z wykorzystaniem formalizmu znanego z teorii grafów. Chęć poznania mechanizmów zachodzących w układach rzeczywistych z różnych obszarów, takich jak biologia, socjologia czy ekonomia, angażuje naukowców zajmujących się różnymi dziedzinami nauki, w tym fizyków. Zrozumienie zachowań i zjawisk zachodzących w badanych układach nie jest możliwe bez stworzenia odpowiednich modeli, które odzwierciedlają własności sieci rzeczywistych. Dyskutowana rozprawa doktorska wpisuje się w ten nurt, proponując model opisu układów złożonych, bliższy rzeczywistej strukturze tychże sieci.

W rozdziale 2. rozprawy zatytułowanym: „Wykładnicze grafy przypadkowe dla sieci ze strukturami blokowymi”, zaprezentowany został model generowania grafów posiadających strukturę blokową, co odpowiada konstrukcji grafu, w którym można wyróżnić dobrze zdefiniowane podgrafy. Ważną cechą wprowadzonego modelu jest możliwość określenia rozkładu stopni wierzchołków,

dzięki czemu generowane grafy mogą odzwierciedlać cechy układów rzeczywistych, takie jak bezskalowy rozkład stopni wierzchołków czy własność małego świata. W podrozdziale 2.2.3. Autor przedstawia wyprowadzenie wzorów na średni stopień wierzchołka w grafie blokowym, dla połączeń wewnątrz danego bloku i pomiędzy blokami, w oparciu o formalizm zespołu statystycznego. W kolejnym podrozdziale zostały zaprezentowane analogiczne obliczenia dla sieci blokowych z korektą na stopnie węzłów, oznaczającą wprowadzenie innego niż jednorodny rozkładu stopni wierzchołków. Poprawka ta jest bardzo istotna, gdyż tworzone z jej uwzględnieniem sieci lepiej oddają strukturę sieci rzeczywistych. Na końcu podrozdziału Autor przedstawia także wzory wyrażające prawdopodobieństwa połączenia dwóch węzłów sieci w przypadku, gdy należą one do tego samego bloku oraz gdy znajdują się w dwóch różnych blokach. Wszystkie obliczenia prowadzone są przy założeniu różnych wielkości bloków tworzących sieć, co także ma uzasadnienie w odniesieniu do struktur rzeczywistych.

Podrozdział 2.3 pracy poświęcony jest przedstawieniu modelu numerycznego, który został użyty do - jak stwierdza Autor - potwierdzenia poprawności przeprowadzonych obliczeń analitycznych. Model ten, działając w oparciu o algorytm Metropolisa-Hastingsa, pozwala na wygenerowanie sieci o zadanych parametrach. Autor prezentuje przykładowe sieci stanowiące realizację klasycznego modelu blokowego oraz modelu z korektą na stopnie wierzchołków. W pierwszym przypadku wynikowa sieć charakteryzuje się jednakowymi wielkościami bloków oraz prawdopodobieństwem połączenia wierzchołków zależnym od przynależności do tego samego lub różnych bloków. Zmodyfikowana wersja modelu prowadzi do możliwości wygenerowania sieci o heterogenicznym rozkładzie stopni wierzchołków, czego wynik zaprezentowany został na rys. 2.2. Z tego, co zostało napisane w rozdziale 2.2.4 rozumiem, że model pozwala także na generowanie sieci o różnych rozmiarach bloków, co jednak nie zostało pokazane.

Kolejny, trzeci rozdział pracy dotyczy zjawiska perkolacji na sieciach, w tym na klasycznych sieciach blokowych (podrozdział 3.3). Podrozdział wstępny zawiera krótkie omówienie klasycznego modelu blokowego, a w podrozdziale 3.2 przedstawiony został formalizm dotyczący badania perkolacji w grafach przypadkowych przy czym materiał w nim zawarty nie stanowi bezpośrednio części analizowanej rozprawy doktorskiej. Przedmiotem rozprawy jest natomiast, przedstawione w podrozdziale 3.3, zjawisko perkolacji w grafach o strukturze blokowej. Autor prezentuje w nim wyprowadzenie wzoru na prawdopodobieństwo przynależności węzła do globalnego, tj. obejmującego różne bloki, klastra perkolacyjnego oraz równanie uwikłane umożliwiające wyznaczenie rozmiaru klastra perkolacyjnego. Dalej, Doktorant prezentuje wyrażenie określające próg perkolacji w klasycznym modelu blokowym. Na rys. 3.3 przedstawiony został diagram fazowy dla perkolacyjnego przejścia fazowego, na którym porównane zostały wyniki otrzymane poprzez obliczenia numeryczne z wynikami analitycznymi, zaś na rys. 3.4 analogiczny wykres, jednak z uwzględnieniem skończonych rozmiarów sieci. Na rysunku 3.5 przedstawione zostało natomiast porównanie wyznaczonego numerycznie i analitycznie rozmiaru globalnego klastra perkolacyjnego w zależności od średniego globalnego stopnia wierzchołka dla sieci o trzech różnych rozmiarach. Zgodnie z zaprezentowanymi wynikami uzyskano dużą zgodność wyników otrzymanych obiema metodami.

Ostatni, nie licząc podsumowania, podrozdział rozdziału 3, stanowi zastosowanie omawianego formalizmu do optymalizacji kosztów w przykładowej, hipotetycznej sieci rzeczywistej o budowie blokowej.

Rozdział 4. rozprawy Pana Bujoka dotyczy analizy własności polskiej i angielskiej sieci semantycznej. Autor prezentuje w nim motywację, stojącą za tworzeniem tego typu sieci oraz omawia pojęcia niezbędne do zrozumienia zależności, w oparciu o które konstruowane są te sieci. Dalej Autor przedstawia porównanie polskiej i angielskiej sieci pod kątem zależności liczby zbiorów synonimów od liczby zawartych w nich leksemów, czyli abstrakcyjnych jednostek systemu języ-

ka, oraz rozkład polisemii dla podstawowych części mowy. W rozdziale 4.4 przedstawiona została analiza grafów utworzonych dla synsetów, tj. zbiorów synonimów, w których krawędzie stanowią relacje hiperonimii czyli nadrzędności, bądź też wyrażają wszystkie możliwe relacje znaczeniowe. Autor stawia tezę, że ze względu na to, iż otrzymany rozkład wielkości klastrów dla sieci z relacjami hiperonimii ma charakter rozkładu o tłustym ogonie, dodanie niewielkiej liczby wiązań do sieci doprowadziłoby do spójności znacznej jej części. W grafie, którego krawędzie odzwierciedlają także inne relacje, zaobserwowano dla obu analizowanych sieci językowych istnienie więcej niż jednego klastra, co oznacza istnienie synsetów, które nie są ze sobą połączone. Doktorant sugeruje, że przyczyna tego faktu tkwi w niepełności danych, chociaż jak sam stwierdza większa liczba takich klastrów w sieci dla języka angielskiego, nad tworzeniem której prace trwają znacznie dłużej, nie do końca potwierdza postawioną tezę. Wydaje mi się, że ciekawe byłoby sprawdzenie znaczeń synsetów tworzących odizolowane klastry.

W dalszej części podrozdziału 4.4 Autor przedstawia analizę omawianych sieci semantycznych w nawiązaniu do wcześniejszych rozdziałów rozprawy. Wskazane zostały ich własności modułowe, wraz z relacjami wprowadzającymi dodatkowo hierarchiczność sieci. Jak zostało stwierdzone na str. 61 ze względu na znaczny stopień skomplikowania sieci nie została podjęta próba ich modelowania.

Podrozdział 4.5 poświęcony jest analizie relacji międzyjęzykowych, które wykorzystywane są m.in. przez programy tłumaczące. Analiza ta wymaga konstrukcji sieci warstwowej, w której poszczególne warstwy odnoszą się do konkretnego języka, natomiast połączenia pomiędzy warstwami wyrażają relacje zachodzące pomiędzy synsetami pochodzącymi z różnych języków, takie jak np. synonimia międzyjęzykowa. Za miarę poprawności połączeń pomiędzy obiema sieciami przyjęto supremację, która to wielkość - zgodnie z informacją umieszczoną na stronie 62 rozprawy - „wyraża liczbę węzłów podporządkowanych danemu węzłowi” i pozwala na określenie ważności węzła w sieci. Otrzymane wyniki pokazują generalnie zgodność znaczeniową pomiędzy dopasowanymi synsetami, przy czym, co wydaje się intuicyjnie słuszne, jest ona lepsza dla pojęć bardziej ogólnych.

Główna moja uwaga krytyczna do omawianej pracy odnosi się do konstrukcji rozdziałów 2 i 3. Pojawiające się we wstępach do obu rozdziałów częściowo powtarzające się informacje o modelu blokowym mogły być przedstawione całościowo raz. Pomimo tego, iż rozdział 2 dotyczy modelowania i analizy sieci blokowych, we wstępie do rozdziału 3 dotyczącego perkolacji w klasycznych grafach przypadkowych pojawia się akapit dotyczący pracy, która jako pierwsza opisywała model blokowy (w rozdziale 2 na str. 22 praca ta jest jedynie wzmiankowana jako jedna z kilku dotyczących badań nad sieciami tego typu). Także dopiero w rozdziale 3 pojawia się rysunek (rys. 3.2 str. 44) przedstawiający ilustrację grafu o strukturze blokowej, wraz z odpowiadającą mu macierzą połączeń. Źródła tego faktu można się dopatrywać w tym, że rozdziały 2-4 stanowią prezentację publikacji, będących podstawą doktoratu. W pracy zabrakło też moim zdaniem komentarza odnośnie możliwości i celowości przeprowadzenia analizy zjawiska perkolacji w sieciach blokowych z korektą na stopnie wierzchołków.

W moim odczuciu, nie ma także spójności pomiędzy zagadnieniami prezentowanymi w rozdziałach 2 i 3 a analizą sieci rzeczywistej, omawianą w rozdziale 4. Interesuje mnie również, dlaczego w pracy analizowana jest polska sieć semantyczna w wersji 2.0 z roku 2009, czyli pierwsza, która została udostępniona, a nie wersja bardziej dojrzała.

Dla porządku odnotowuję także pojawiające się w tekście błędy edytorskie takie jak literówki, błędny odnośnik do rysunku (str. 57) czy niespójne użycie formy liczby mnogiej wyrazu „klaster” (w tekście najpierw pojawia się forma ”klasterów”, później „klastrów”).

Pomimo powyższych uwag, w moim odczuciu dorobek prezentowany w niniejszej pracy stanowi interesujący i cenny wkład w badania nad układami złożonymi. Może też stanowić inspirację dla

dalszych badań.

Podsumowując, uważam, że omawiana rozprawa doktorska spełnia kryteria określone w ustawie o *stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki*, dlatego też wnoszę o dopuszczenie mgr. inż. Maksymiliana Bujoka do dalszych etapów przewodu doktorskiego.

MKrawczyk