

Warszawa, 20 marca 2023 r.

prof. dr hab. Zbigniew Lonc
Wydział Matematyki i Nauk Informatycznych
Politechnika Warszawska
ul. Koszykowa 75
00-662 Warszawa

**Recenzja w postępowaniu habilitacyjnym
doktora inżyniera Krzysztofa Turowskiego**

Dr Krzysztof Turowski przedstawił jako swoje osiągnięcie naukowe, wymagane w art. 219 ust. 1 pkt 2b ustawy Prawo o szkolnictwie wyższym i nauce, cykl następujących, powiązanych tematycznie, pięciu artykułów naukowych:

1. Krzysztof Turowski, Wojciech Szpankowski, Towards Degree Distribution of a Duplication-Divergence Graph Model, *Electronic Journal of Combinatorics*, 28(1) (2021), P1.18.
2. Alan Frieze, Krzysztof Turowski, Wojciech Szpankowski, Degree Distribution for Duplication-Divergence Graphs: Large Deviations, *46th International Workshop on Graph-Theoretic Concepts in Computer Science, WG 2020*, Leeds, UK, June 24-26, 2020. Lecture Notes in Computer Science 12301, s. 226-237.
3. Alan Frieze, Krzysztof Turowski, Wojciech Szpankowski, The concentration of the maximum degree in the duplication-divergence models, *Proceedings of 27th International Conference of Computing and Combinatorics, COCOON 2021*, Tainan, Taiwan, October 24-26, 2021. Lecture Notes in Computer Science 13025, s. 413-424.
4. Philippe Jacquet, Krzysztof Turowski, Wojciech Szpankowski, Power-Law Degree Distribution in the Connected Component of a Duplication Graph, *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, AofA 2020*, June 15-19, 2020, Klagenfurt, Austria (Virtual Conference). LIPIcs 159, s. 16:1-16:14.

5. Krzysztof Turowski, Abram Magner, Wojciech Szpankowski, Compression of Dynamic Graphs Generated by a Duplication Model, *Algorithmica* 82(9) (2020), s. 2687-2707.

To osiągnięcie naukowe zatytułowane jest „Analiza strukturalna i kompresja dla duplikacyjnych modeli grafów losowych”. Powyższe prace ukazały się w czasopiśmie i recenzowanych materiałach z konferencji międzynarodowych znajdujących się w odpowiednim wykazie ministerialnym.

Wszystkie prace składając się na to osiągnięcie są współautorskie, pierwsza z jednym, a pozostałe z dwoma współautorami. Ze znajdujących się w dokumentacji oświadczeń Habilitanta i jego współautorów wynika, że we wszystkich pięciu pracach dr Turowski jest wiodącym współautorem. Jego przybliżony wkład w powstanie tych prac wynosi kolejno: 90%, 50%, 55%, 65% i 65%.

Przedmiotem badań Habilitanta, których wyniki znajdują się w powyższych pracach jest probabilistyczne zachowanie modelu sieci duplikacyjnej zaproponowanego przez Solégo i Pastor-Satorrasa w 2002 r. W modelu tym nowy wierzchołek przybywający w chwili t najpierw losowo wybiera istniejący wierzchołek i łączy się krawędzią z jego sąsiadami z prawdopodobieństwem p , a następnie łączy się z innymi wierzchołkami z prawdopodobieństwem r/t , gdzie p i r to ustalone parametry w tym modelu.

Motywacją rozważania takiego modelu sieci, w którym nowe wierzchołki pojawiają się w wyniku odpowiednio zdefiniowanego mechanizmu duplikacji i mutacji, były obserwacje zjawisk zachodzących w pewnych sieciach biologicznych. Podobnymi mechanizmami wydają się również rządzić niektóre sieci społeczne. Badanie tego modelu ograniczało się dotąd tylko do przeprowadzenia symulacji komputerowych oraz obliczeń przybliżeń. W pełni rygorystyczne, systematyczne badania własności tego modelu nie były dotąd wykonane. Tę lukę wypełniają w pewnym stopniu wyniki zawarte w ocenianym osiągnięciu naukowym.

W pracy [1] zbadane zostało zachowanie średniego stopnia i stopnia ustalonego wierzchołka w grafie losowym wygenerowanym przez model Solégo i Pastor-Satorrasa. Znaleziono zostały dokładne asymptotyki dla wartości oczekiwanych i wariancji średniego stopnia oraz stopnia ustalonego wierzchołka, jako funkcji liczby wierzchołków. Aby uzyskać te wyniki Autorzy najpierw znaleźli rekurencyjne równania dla tych funkcji, a następnie opracowali metodę, która stosuje się nie tylko do rozwiązania tych równań, ale również równań ogólniejszych podobnego typu. Jest to technicznie skompli-

kowana metoda, która wymaga pokazania pewnych własności funkcji gamma Eulera oraz użycia narzędzi z zakresu teorii funkcji hipergeometrycznych.

W pracy [2] Autorzy kontynuują linię badań podjętą w pracy [1] i dowodzą twierdzeń o koncentracji rozkładu dla tych samych zmiennych losowych, tj. średniego stopnia i stopnia ustalonego wierzchołka, wokół ich wartości oczekiwanych. W tym celu odpowiednio szacują ogony rozkładów tych zmiennych. Okazuje się, że wartość średniego stopnia z prawdopodobieństwem $1 - O(t^{-A})$, gdzie $A > 0$ to dowolna stała, odbiega od swojej wartości oczekiwanej co najwyżej o czynnik polilogarytmiczny. Podobny wynik zachodzi dla stopni ustalonych wierzchołków. W technicznie skomplikowanych dowodach, oprócz wyników z pracy [1], Autorzy wykorzystują funkcje tworzące momenty oraz twierdzenie Chernoffa.

Artykuł [3] zawiera podobny wynik jak w pracy [3] tylko dotyczący maksymalnego stopnia grafu wygenerowanego przez model Solégo i Pastor-Satorrasa. Dla wartości parametru $\frac{1}{2} < p < 1$ pokazują, że maksymalny stopień jest asymptotycznie skoncentrowany wokół t^p , tj. odbiega od tej wartości co najwyżej o czynnik polilogarytmiczny. Wynik ten został uzyskany podobnymi metodami jak w pracy [2], choć dowód wymagał też nowych pomysłów.

Praca [4] dotyczy tzw. modelu czystej duplikacji zaproponowanego w 2003 r. przez Chung i in., gdzie zakłada się, że nowy wierzchołek przybywający w chwili t jedynie losowo wybiera istniejący wierzchołek i łączy się krawędzią z jego sąsiadami z prawdopodobieństwem p . Formalnie, jest to szczególnie przypadek modelu Solégo i Pastor-Satorrasa, gdy $r = 0$. Sieć nazywamy bezskalową jeśli odsetek wierzchołków stopnia k dąży do ustalonej wartości proporcjonalnej do $k^{-\gamma}$ dla pewnego ustalonego parametru γ . Własność bezskalowości jest ważna ze względu na to, że wiele sieci biologicznych i społecznych posiada tę własność. W pracy [4] Autorzy pokazali piękny wynik, że dla wartości parametru p z przedziału $(0, \frac{1}{e})$ model czystej duplikacji generuje grafy bezskalowe ze współczynnikiem γ z przedziału $(2, 3)$ (dokładniej, ich spójna składowa ma tę własność). Autorzy podali zarówno wartość współczynnika γ jak i znaleźli asymptotykę dla współczynnika proporcjonalności odsetka wierzchołków stopnia k do $k^{-\gamma}$ dla $p \in (0, \frac{1}{e})$. Wynik pracy [4] stanowi potwierdzenie wcześniejszego przypuszczenia innych autorów. Swoje rezultaty Autorzy pokazali metodami analizy zespolonej z wykorzystaniem transformaty Mellina oraz metody residuów.

W pracy [5] rozważany jest jeszcze bardziej szczególny przypadek modelu Solégo i Pastor-Satorrasa, który nazywa się modelem pełnej duplikacji. W tym modelu nowy wierzchołek losowo wybiera istniejący wierzchołek i staje

się jego kopią (ma tych samych sąsiadów). Autorów interesuje problem ile bitów potrzeba do reprezentacji grafów wygenerowanych przez ten model w wersjach dla grafów z etykietami na wierzchołkach i bez etykiet. Główne wyniki tej pracy to dokładne oszacowania asymptotycznego tempa wzrostu entropii w obu przypadkach oraz znalezienie bliskich optymalnym algorytmów kompresji w tych przypadkach. W dowodzie teoretycznych rezultatów pracy korzysta się głównie z metod analitycznych oraz z wcześniejszego wyniku Łuczka i in. na temat związku entropii grafów z etykietami na wierzchołkach i bez etykiet wygenerowanymi przez pewne modele grafów losowych.

W moim przekonaniu rezultaty zawarte w pięciu omówionych wyżej pracach składających się na osiągnięcie naukowe dr. inż. Krzysztofa Turowskiego stanowią znaczny i istotny wkład w poznanie dość słabo do tej pory zbadanych własności duplikacyjnych modeli grafów losowych. Uważam, że ta tematyka badawcza jest ważna i dobrze umotywowana przez fakt, że rozważane modele powstały na podstawie obserwacji rzeczywistych sieci biologicznych i społecznych. Rezultaty prac [1], [2] i [3] stanowią znaczący postęp w zrozumieniu ogólnej struktury grafów generowanych przez model Solégo i Pastor-Satorrasa, w szczególności rozkładu stopni wierzchołków. Praca [4] zawiera wynik, który jest odpowiedzią na wcześniej otwarty problem i dowodzi własności bezskalowości, co jest argumentem dla tezy, że rozważany model dobrze przybliża pewnego rodzaju sieci realnie istniejące. W końcu praca [5] zawiera algorytmy pozwalające w optymalny sposób kompresować grafy losowe w rozważanym modelu, co może mieć znaczenie praktyczne związane z przechowywaniem i przetwarzaniem wielkoskalowych danych grafowych. Tak więc bardzo dobrze oceniam poziom i znaczenie rezultatów zawartych w osiągnięciu naukowym dr. Turowskiego. Podobnie wysoko oceniam warsztat badawczy Habilitanta. Dowody twierdzeń zawartych w omówionych pracach używają różnorodnych narzędzi zarówno probabilistycznych, analitycznych, jak i kombinatorycznych. Widać dużą biegłość w posługiwaniu się tymi metodami.

Uważam więc, że rezultaty uzyskane w cyklu omówionych pięciu prac dr. Turowskiego stanowią znaczny wkład w rozwój dyscypliny informatyka, a zatem spełniony jest warunek sformułowany w art. 219 ust. 1 pkt 2b ustawy Prawo o szkolnictwie wyższym i nauce.

Pozostały dorobek publikacyjny dr. inż. Krzysztofa Turowskiego składa się 17 artykułów naukowych (nie licząc wersji konferencyjnych prac opublikowanych później w czasopiśmie).

Dwie z tych prac również dotyczą modeli duplikacyjnych. W pracy [C1]¹ podjęto problem dopasowania parametrów modelu duplikacyjnego tak, aby liczba automorfizmów generowanych przez ten model grafów była bliska tej liczbie w rzeczywistych sieciach. Druga praca z tej grupy ([C2]) dotyczy tzw. problemu archeologii w sieci, tj. rekonstrukcji ewolucji sieci w czasie. W dwóch następnych pracach ([B1] i [B2]) Autorzy badają zagadnienia kompresji i entropii pewnego rodzaju drzew losowych. Publikacje z tej grupy dotyczą zagadnień, które są przedmiotem badań dużej grupy badaczy o czym świadczą dość liczne cytowania.

Praca [D1] jest ubocznym produktem badań, których rezultaty są zawarte w pracy [5]. Podano w niej precyzyjną asymptotykę funkcji entropii dla zmiennej o rozkładzie wielomianowym Dirichleta.

Na uwagę zasługuje praca [E1], gdzie rozważano pewną klasę problemów szeregowania zadań z ograniczeniami dotyczącymi wyboru maszyn, które mogą realizować zadania. Podano szereg algorytmów dokładnych i przybliżonych oraz wyników złożonościowych dla różnych wariantów tego problemu. Jest to w moim przekonaniu bardzo solidna praca.

Seria następnych ośmiu prac Habilitanta dotyczy pewnych kolorowań grafów, zwanych kolorowaniami szkieletowymi, mających związek z problemem przydziału częstotliwości. Ta tematyka była przedmiotem badań w ramach pracy doktorskiej dr. Turowskiego i jest nadal przez niego kontynuowana. Z pozostałych trzech prac dwie dotyczą gier kombinatorycznych, a jedna problemu rekonstrukcji hipergrafów.

Ten krótki przegląd tematyki badawczej i osiągnięć dr. Turowskiego pokazuje, że nie ogranicza się on do jednego nurtu badań, lecz podejmuje dość różnorodną tematykę, w ramach której udaje mu się uzyskać istotne rezultaty. Wyniki tych badań są opublikowane w dobrych czasopismach naukowych.

Co prawda dane naukowometryczne dr. Turowskiego nie są może zbyt imponujące, ale najlepsze, w moim przekonaniu, jego prace (w tym te wchodzące w zakres osiągnięcia habilitacyjnego) powstały stosunkowo niedawno więc sądzę, że liczba cytowań wkrótce znacząco wzrośnie.

Innym przejawem aktywności naukowej dr. Turowskiego są wystąpienia na konferencjach, których jak dotąd było 10, z czego 8 po uzyskaniu stopnia naukowego doktora.

¹W niniejszej recenzji stosuję numerację prac zgodną ze znajdującym się w dokumentacji Autoreferatem Habilitanta.

Ważną rolę w rozwoju naukowym Habilitanta odegrały dwa długoterminowe staże poddoktorskie, które odbył w Purdue University w USA. Pracował tam w grupie naukowej prof. Wojciecha Szpankowskiego. Dr Turowski był także beneficjentem grantu Sonata przyznanego przez NCN. Ponadto był wykonawcą trzech grantów przyznanych przez instytucje krajowe oraz jednego przyznanego przez National Science Foundation w USA. Obecnie jest wykonawcą kolejnego grantu przyznanego przez NSF w USA.

Wszystkie te informacje świadczą o dużej aktywności naukowej dr. Turowskiego. Prawie wszystkie prace dr. Turowskiego są współautorskie. Większość z nich powstała we współpracy z zespołami z Purdue University i Politechniki Gdańskiej. Dlatego nie mam wątpliwości, że warunek sformułowany w art. 219 ust. 1 pkt 3 Ustawy mówiący o wykazywaniu się kandydata do stopnia doktora habilitowanego „istotną aktywnością naukową (...) realizowaną w więcej niż jednej uczelni, instytucji naukowej (...), w szczególności zagranicznej” jest spełniony.

Konkludując, w mojej opinii, dr inż. Krzysztof Turowski spełnia ustawowe i zwyczajowe wymogi uzyskania stopnia naukowego doktora habilitowanego w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie informatyka.