

Łódź, 14 czerwca 2023 r.

dr hab. inż. Agnieszka Wosiak, prof. uczelni,
Instytut Informatyki
Wydział Fizyki Technicznej, Informatyki i Matematyki Stosowanej
Politechnika Łódzka
Al. Politechniki 8
93-590 Łódź

RECENZJA

dotycząca wniosku w postępowaniu o nadanie stopnia doktora habilitowanego
w dziedzinie nauk inżynieryjno-technicznych
w dyscyplinie informatyka techniczna i telekomunikacja

Pana dra inż. Marcina Michalaka

PODSTAWA OPRACOWANIA RECENZJI

Podstawę opracowania niniejszej recenzji stanowi Uchwała nr 39/2023 Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej z dnia 28 marca 2023 r. powołująca mnie w skład komisji habilitacyjnej w postępowaniu w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja Panu dr inż. Marcinowi Michalakowi i powierzająca mi funkcję recenzenta w tym postępowaniu.

Podstawę prawną w zakresie kryteriów branych pod uwagę przy ocenie osiągnięć naukowych Kandydata ubiegającego się o stopień doktora habilitowanego jest art. 219 ust. 1 pkt 2 Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce.

Recenzja została przygotowana na podstawie przekazanej mi w formie elektronicznej dokumentacji przygotowanej przez Kandydata, która obejmowała:

- kopię wniosku o przeprowadzenie postępowania habilitacyjnego,
- dane wnioskodawcy,
- autoreferat,
- wykaz osiągnięć naukowych,
- kopie publikacji naukowych.

SYLWETKA KANDYDATA

Pan dr inż. Marcin Michalak ukończył z wynikiem bardzo dobrym Wydział Automatyki, Elektroniki i Informatyki Politechniki Śląskiej w 2005 roku uzyskując tytuł zawodowy magistra inżyniera informatyki.

W latach 2005 – 2009 był doktorantem w Instytucie Informatyki Politechniki Śląskiej. W 2009 roku na Wydziale Automatyki, Elektroniki i Informatyki Politechniki Śląskiej obronił rozprawę doktorską zatytułowaną „Estymacja funkcji regresji w predykcji szeregów czasowych” uzyskując stopień naukowy doktora nauk technicznych w dyscyplinie informatyka.

Od 2009 do 2012 roku Kandydat był zatrudniony w Głównym Instytucie Górnictwa, najpierw jako inżynier, a od kwietnia 2010 roku na stanowisku adiunkta.

Równolegle, od 2009 roku Pan dr inż. Marcin Michalak związany jest zawodowo z Wydziałem Automatyki, Elektroniki i Informatyki Politechniki Śląskiej. Początkowo, do roku 2014 pracował na stanowisku asystenta w Instytucie Informatyki Politechniki Śląskiej. Od 2014 roku zatrudniony jest na stanowisku adiunkta w Katedrze Sieci i Systemów Komputerowych (wcześniej Instytut Informatyki).

Od 2016 roku pracuje również jako starszy specjalista w ramach Sieci Badawczej Łukasiewicz w Instytucie Technik Innowacyjnych EMAG (wcześniej Instytut Technik Innowacyjnych EMAG) w Katowicach.

OCENA OSIĄGNIĘCIA NAUKOWEGO

Jako osiągnięcie naukowe stanowiące podstawę do uzyskania stopnia doktora habilitowanego Kandydat wskazał cykl dziesięciu powiązanych tematycznie publikacji pod wspólnym tytułem „Wnioskowanie boolowskie w biklasteryzacji”.

W skład cyklu wchodzi poniżej wymienione pozycje. Przy każdej pozycji podaję deklarowany przez Kandydata wkład w ich powstanie – dołączona do wniosku dokumentacja elektroniczna nie zawierała oświadczeń współautorów.

- [1] M. Michalak (2022): “Theoretical Backgrounds of Boolean Reasoning Based Binary n-clustering”, Knowledge and Information Systems, vol. 64, str. 2171-2188 (IF = 2.531 (2021); MEiN = 100 pkt.; udział = 100%)
- [2] M. Michalak (2021): “Hierarchical heuristics for Boolean-reasoning-based binary bicluster induction”, Acta Informatica, vol. 59(6), str. 673-685 (IF = 0.871 (2021); MEiN = 70 pkt.; udział = 100%)
- [3] M. Michalak, R. Jaksik, D. Ślęzak (2020): “Heuristic Search of Exact Biclusters in Binary Data”, International Journal of Applied Mathematics and Computer Science, vol. 30(1), str. 161-171, (IF = 1.417; MEiN = 100 pkt.; udział = 80%)
- [4] M. Michalak, D. Ślęzak (2019): “On Boolean Representation of Continuous Data Biclustering”, Fundamenta Informaticae, vol. 167(3), str. 193-217, 2019 (IF = 1.298; MEiN = 70 pkt.; udział = 80%)

- [5] M. Michalak, D. Ślęzak (2018): "Boolean Representation for Exact Biclustering", *Fundamenta Informaticae*, vol. 161(3), str. 275-297 (IF = 1.204; MEiN = 70 pkt.; udział = 80%)
- [6] M. Michalak (2019): "Induction of Centre-Based Biclusters in Terms of Boolean Reasoning", w *Man-Machine Interactions 6. ICMMI 2019*, str. A. Gruca et al., 239-248, vol. 1061, Springer, Cham. (MEiN = 20 pkt.; udział = 100%)
- [7] M. Michalak, M. Stawarz (2011): "Generating and Postprocessing of Biclusters from Discrete Value Matrices", w: *Computational Collective Intelligence. Technologies and Applications. ICCCI 2011*, Jędrzejowicz P. et al., Eds., str. 103-112, vol. 6922, Springer, Berlin, Heidelberg, (WoS; MEiN=15 pkt., udział = 70%)
- [8] M. Michalak (2022): "Boolean reasoning in biclustering", w *Artificial Intelligence and Data Processing. The monograph presenting the achievements of the Silesian University of Technology research staff.*, P. Kasprowski et al., Eds., str. 185-190, Politechnika Śląska, (MEiN = 20 pkt., udział = 100%)
- [9] M. Michalak, R. Jaksik (2022): "Application of boolean reasoning paradigm in biomedical data biclustering", w *Artificial Intelligence and Data Processing. The monograph presenting the achievements of the Silesian University of Technology research staff.*, P. Kasprowski et al., Eds., str. 319-323, Politechnika Śląska, (MEiN =20 pkt., udział = 75%)
- [10] M. Michalak (2019): „Boolean Biclustering Review and Perspectives”, w *PP-RAI'2019 Polskie Porozumienie na Rzecz Rozwoju Sztucznej Inteligencji*, 16-18.10.2019, Wrocław, Poland: Conference Proceedings, I. Czarnowski et al., Eds., str. 265-268, Politechnika Wroclawska, (MEiN=20 pkt., udział = 100%)

Wśród wymienionych pozycji pięć publikacji wydanych zostało w czasopismach indeksowanych w bazie *Journal of Citation Reports*), dwie stanowią materiały konferencyjne, zaś trzy ostatnie to rozdziały w opracowaniach wieloautorskich (monografiach). Warto zauważyć, że w pięciu publikacjach cyklu Kandydat jest jedynym autorem, zaś w pozostałych pięciu jego udział jest znaczący i waha się od 70% do 80%.

Sumaryczna liczba punktów MEiN dla publikacji w cyklu wynosi 490 punktów, zaś wskaźnik wpływu (ang. *Impact Factor*) jest równy 7.321. Po uwzględnieniu wag wynikających z udziału współautorskiego wskaźniki te wnoszą odpowiednio 362 punkty MEiN oraz IF=6.535.

Najważniejszym oryginalnym osiągnięciem przedstawionym w każdej pracy cyklu było wykazanie, że poszukiwanie wzorców biklasteryzacyjnych w danych można przedstawić równoważnie jako poszukiwanie implikantów funkcji boolowskiej, odpowiadających poszukiwanym wzorcom.

W ogólnym ujęciu biklasteryzacja stosowana jest w analizie danych do jednoczesnego grupowania obiektów i cech. W kontekście macierzy danych, biklasteryzacja polega na poszukiwaniu podmacierzy, które są jak najbardziej spójne pod względem wartości. Biklasteryzacja znajduje wiele zastosowań i jest szczególnie istotna w analizie danych o wysokiej wymiarowości, takich jak dane genetyczne, dane z mikromacierzy związane z ekspresją genów, ale również w analizie danych tekstowych i sieci społecznościowych, w tym również w kontekście systemów rekomendacyjnych.

Pracą otwierającą publikacje z cyklu wskazanego przez Kandydata jest praca [7]. Opisane w niej zostały pierwsze próby wykorzystania funkcji boolowskich w zagadnieniu biklasteryzacji. W szczególności opisane zostało podejście polegające na generowaniu dwóch macierzy odróżnialności: jednej dla wierszy i drugiej dla kolumn, przy czym w komórce (i, j) macierzy odróżnialności dla kolumn znajdowały się zmienne logiczne odpowiadające tym wierszom, dla których wartość w kolumnie i stanowiło 0 a w kolumnie j stanowiło 1. W oparciu o macierz odróżnialności w kolejnym kroku budowana była funkcja boolowska jako koniunkcja alternatyw z każdej komórki tej macierzy. Literały tworzące implikant prosty odpowiadały tzw. półbiklastrowi wierszy albo kolumn. W przedostatnim kroku wszystkie półbiklastrowy wierszy był łączone ze wszystkimi półbiklastrowymi kolumn, a w ostatnim natomiast kroku usuwano biklastrowy puste. Nowością podejścia było odkrywanie biklastrowy przez analizę funkcji boolowskich. Opisana metoda miała jednak ograniczenia polegające na konieczności analizowania dwóch takich funkcji oraz końcowym ponownym przetworzeniu uzyskanych rezultatów. Ponadto, metoda nie zapewniała maksymalności w sensie inkluzji uzyskanych biklastrowy, ani dokładności (biklastrowy mógł zawierać zera). Te ograniczenia stanowiły podstawy do dalszych badań Habilitanta, opisanych w kolejnych pracach cyklu.

Habilitant na podstawie opracowanych twierdzeń i ich dowodów przedstawionych w pracach z cyklu wykazał możliwość ich zastosowania w skutecznym poszukiwaniu:

- opisanych w pozycji [1] n -klastrowy jedynek lub zer w hiperkostkach binarnych: poszukiwanie wzorców odbywa się za pomocą paradygmatu wnioskowania boolowskiego, gdzie oryginalne dane są kodowane do formuły boolowskiej, a jej najważniejsze implikanty są równoważne z rozwiązaniami pierwotnego problemu; artykuł przedstawia zastosowanie n -klastrowy bazujących na wnioskowaniu boolowskim dla niewielkich, sztucznie wygenerowanych danych trzy- i czterowymiarowych, a także dla wybranego zbioru danych biomedycznych pobranego z repozytorium TCGA (The Cancer Genome Atlas),

- opisanych w pozycji [5] biklastrów dokładnych w macierzach dyskretnych oraz biklastrów jedynek lub zer w macierzach binarnych z dodatkowym uwzględnieniem potraktowania macierzy dyskretnych jako obrazów i zastosowania technik przetwarzania obrazów w celu zmniejszenia złożoności obliczeniowej procesu analizy,
- opisanych w pozycji [4] biklastrów podobieństwa oraz biklastrów chaosu w macierzach ciągłych: za główny wkład tej pracy należy wskazać matematycznie udowodnioną boolowską reprezentację biklastrów dla danych ciągłych, drugim, istotnym wkładem pracy jest wprowadzenie pojęcia odwrotności biklastrów podobieństwa określanego jako biklastry chaosu, czyli takich skupień, w których wszystkie elementy różnią się od siebie na poziomie przekraczającym określoną wartość progową.

Ponadto:

- w pracy [3] zaproponowane zostały: modyfikacja strategii Johnsona poszukiwania implikantów funkcji boolowskiej, zapobiegająca znajdowaniu biklastrów pustych oraz strategia sekwencyjnego poszukiwania biklastrów w macierzy binarnej, w celu pokrycia biklastrami wszystkich „jedynek” w macierzy,
- w pracy [2] Habilitant kontynuując badania opisane w pracy [3] zaproponował hierarchiczną heurystykę poszukiwań biklastrów, która pozwala na generowanie większej liczby ogólnych wzorców, które z kolei mogą się nakładać; przeprowadzone badania eksperymentalne wykazały zdolność do odnajdywania rozległych biklastrów, pokrywających znaczący obszar danych.

Kolejne prace cyklu [8], [9] oraz [10] stanowią zbiorcze zestawienia powyższych osiągnięć zaprezentowane odpowiednio na konferencji krajowej ([8]) oraz w pracy zbiorowej opisującej działalność naukową pracowników Politechniki Śląskiej, co można uznać za słabszą stronę recenzowanego cyklu.

Warto jednak podkreślić perspektywiczny rozwój badań prowadzonych przez Kandydata, w szczególności w kontekście analizy danych masowych i aspektów związanych ze zrównolegleniem obliczeń, które to tematy już na obecnym etapie wskazywane są przez niego w ramach kolejnych prac będących w przygotowaniu.

Biorąc pod uwagę jakość uzyskanych wyników, opublikowanych w międzynarodowych czasopiśmie o szerokim zasięgu wpływu oraz znaczący udział habilitanta w ich powstaniu, można stwierdzić, że **osiągnięcie**, polegające na dostarczeniu twierdzeń i ich dowodów w celu wykazania, że poszukiwanie wzorców biklasteryzacyjnych w danych można przedstawić równoważnie jako poszukiwanie implikantów funkcji boolowskiej, **stanowi istotny wkład w rozwój wiedzy z zakresu dyscypliny informatyka techniczna i telekomunikacja.**

OCENA AKTYWNOŚCI NAUKOWO-BADAWCZEJ

Pan dr Marcin Michalak jest osobą aktywną naukowo, zarówno w Polsce jak i na świecie. W dorobku naukowym Kandydata, oprócz publikacji wskazanych w cyklu związanym z osiągnięciem naukowym, znajduje się osiem publikacji w czasopismach znajdujących się w bazie Journal Citation Reports (JCR), 13 opublikowanych prac naukowych spoza bazy JCR oraz 33 prace przygotowane w ramach opracowań zbiorowych i dokumentacji prac badawczych. Tematyka opracowanych publikacji wskazuje na szerokie zainteresowania naukowe Kandydata, obejmujące między innymi:

- biklasteryzację danych w ujęciu szerszym niż rozważany w ramach cyklu paradygmat wnioskowania boolowskiego, w szczególności jej zastosowanie w zagadnieniach związanych z monitorowaniem pracy urzędów górniczych,
- aspekty związane z indukcją reguł decyzyjnych i oceną ich jakości na poszczególnych etapach ich generowania i stosowania, jak również z odniesieniem się do zagadnień nie tylko klasyfikacyjnych, ale również związanych z regresją i analizą przeżycia,
- diagnostykę urzędów i opracowywanie modeli diagnostycznych, które to prace powiązane są z intensywnym udziałem Habilitanta w projektach badawczych,
- analizę szeregów czasowych oraz wybrane aspekty analizy danych, w tym wykrywanie obserwacji odstających w zbiorach danych powiązanych z prognozowaniem warunków środowiskowych w kopalni węgla kamiennego oraz tematyką bezpieczeństwa w ruchu sieciowym,
- analizę danych wielospektralnych w zagadnieniach dotyczących chorób nowotworowych skóry, ale również wykrywania wybranych fragmentów siatkówki oka.

Udział Habilitanta w opracowaniu prac waha się w przedziale od 15% - 60% przy publikacjach wieloautorskich do 100% w przypadku wyłącznego autorstwa Habilitanta. Sumaryczny współczynnik wpływu według listy Journal Citation Reports zgodnie z rokiem opublikowania wynosi 23.412, zaś po uwzględnieniu współczynników wagowych wynikających z udziału w opracowaniu prac wynosi 12.334. Wyniki badań prowadzonych przez Habilitanta są zauważalne w międzynarodowym środowisku, o czym świadczy wysoka i rosnąca liczba cytowań publikacji. Ich liczba według bazy Web of Science na dzień składania wniosku wynosiła 136, w tym 119 bez autocytowań. Należy zwrócić uwagę, że od tego czasu liczby te odpowiednio wzrosły i wynoszą obecnie odpowiednio 187 i 143. Indeks Hirscha według bazy Web Of Science wynosi 7.

Habilitant od czasu uzyskania stopnia doktora uczestniczył w 32 konferencjach tematycznych, przy czym w 21 w charakterze prelegenta wygłaszającego referat, zaś w 11 prezentacja wyników badań miała charakter plakatu. Pracował również w komitetach organizacyjnych międzynarodowych

konferencji: ICT Innovations, International Joint Conference on Rough Sets oraz International Conference on Enterprise Information Systems.

Ponadto, Habilitant odbył dwa staże zagraniczne w Universidad Pablo de Olavide, Sewilla, Hiszpania, jeden w latach 2020/2021 oraz drugi w 2022 roku. Miał również miesięczny staż naukowy z Instytucie Technik Innowacyjnych EMAG w Katowicach.

Kandydat pełnił rolę redaktora zarządzającego (ang. Managing Editor) w czasopiśmie „Mining – Informatics, Automation and Electrical Engineering”, zaś obecnie pełni rolę redaktora gościnnego (ang. Guest Editor) w numerze specjalnym czasopisma Symmetry. Ponadto, na zaproszenie komitetów naukowych periodyków, był recenzentem 26 artykułów w uznanych czasopismach, między innymi w Neural Networks, Computational Intelligence, Biomedical Signal Processing and Control I International Journal of Applied Mathematics and Computer Science.

Na uwagę zasługuje duża aktywność Kandydata w obszarze związanym z udziałem w projektach badawczych o zasięgu międzynarodowym oraz krajowym. Habilitant ma w swoim dorobku 25 takich projektów. W roli wykonawcy uczestniczył w projektach naukowych finansowanych w ramach różnych programów Narodowego Centrum Nauki oraz Narodowego Centrum Badań i Rozwoju. Ma również doświadczenie w kierowaniu pracami w ramach projektów badawczych, jak na przykład projektem „Budowa narzędzia dedykowanego dla ochrony użytkowników, systemów i urządzeń Internetu Rzeczy, w oparciu o uczenie maszynowe i analizę behawioralną” czy „Analiza warunków i okoliczności występowania oraz analiza zależności czasowych i przyczynowo – skutkowych w wypadkach w kopalniach węgla kamiennego”. Ponadto współpracuje lub współpracował z pracownikami innych ośrodków naukowych w kraju i za granicą, w szczególności z Universidad Pablo de Olavide, Sewilla, Hiszpania oraz Głównym Instytutem Górnictwa. Współpraca ta owocuje projektami naukowymi, czego przykładem może być projekt pod nazwą „Kernel density estimation-based data analytics” finansowany przez hiszpańskie Ministerstwo Nauki i Innowacji (hiszp. Ministerio de Ciencia e Innovación). Za pracę naukową, w tym w zespołach badawczych, Kandydat otrzymał kilka nagród, choć ich powiązanie z osiągnięciem naukowym jest pośrednie.

Reasumując, aktywność naukowa Habilitanta jest wielopłaszczyznowa i – w szczególności z perspektywy współpracy międzynarodowej – na dobrym poziomie. Powyższe rodzaje prac wskazują na pełen zakres kompetencji potrzebnych do samodzielnej pracy po ewentualnym otrzymaniu stopnia doktora habilitowanego. Ponadto potwierdzają umiejętność pracy zespołowej.

OCENA DOROBKU DYDAKTYCZNEGO, ORGANIZACYJNEGO I POPULARYZATORSKIEGO

Pan dr Marcin Michalak aktywnie uczestniczy w realizacji procesu dydaktycznego w Politechnice Śląskiej, choć aktywność ta jest prowadzona z nieco mniejszą intensywnością niż badawcza, co można jednak wytłumaczyć dużym zaangażowaniem Kandydata w projekty o charakterze naukowym.

Habilitant prowadzi szereg różnorodnych kursów przede wszystkim z zakresu analizy danych w języku polskim i w języku angielskim dla studentów kierunków: Informatyka, Informatics (studia w języku angielskim), Control, Electronics and Information Engineering oraz Matematyka. Ponadto był opiekunem pięciu prac dyplomowych magisterskich oraz dwóch prac inżynierskich.

Z przedstawionej dokumentacji nie wynika niestety, jak umiejętności dydaktyczne Kandydata są oceniane przez studentów i czy takie oceny były przeprowadzane.

Kandydat pełni lub pełnił funkcję promotora pomocniczego w dwóch już zamkniętych postępowaniach doktorskich (obrony 2014 roku) oraz w dwóch postępowaniach w dalszym ciągu otwartych (otwarcie w 2017 roku oraz w 2020). Ponadto w jednym postępowaniu doktorskim sprawował funkcję opiekuna naukowego ze strony pracodawcy. W latach 2009-2015 był również opiekunem naukowym doktorantów pracujących nad projektem związanym z utworzeniem nowej specjalności Eksploracja Danych (ang. *Data Mining*) w ramach studiów doktoranckich.

Podsumowując, działalność dydaktyczna, organizacyjna i popularyzatorska Habilitanta ma zadowalający zakres i jest wystarczająca.

WNIOSEK KOŃCOWY

Stwierdzam, że Pan dr inż. Marcin Michalak spełnia wymogi ustawowe dotyczące stopnia naukowego doktora habilitowanego.

Uznaję, że przedłożone do oceny osiągnięcie naukowe pt. „Wnioskowanie boolowskie w biklasteryzacji” wnosi znaczny wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja.

Również pozostałą działalność naukowo-badawczą, dydaktyczną i organizacyjną oceniam dobrze.

Podsumowując, pozytywnie oceniam wniosek Kandydata i wnoszę o kontynuowanie postępowania w sprawie nadania dr inż. Marcinowi Michalakowi stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.