

RPIT-MPI. 24.07.2023

Warszawa, dnia 21 lipca 2023 r.

Prof. dr hab. inż. Andrzej Skowron
Instytut Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa
skowron@mimuw.edu.pl

**Recenzja osiągnięcia naukowego
„Wnioskowanie boolowskie w biklasteryzacji”
oraz całości dorobku naukowego, dydaktycznego i organizacyjnego
dr. inż. Marcina Michalaka
w związku z postępowaniem o nadanie mu stopnia doktora habilitowanego
w dziedzinie nauk inżynieryjno-technicznych
w dyscyplinie informatyka techniczna i telekomunikacja**

Niniejsza recenzja została przygotowana w odpowiedzi na pismo z dnia 28 marca 2023 roku, Prof. dr. hab. inż. Andrzeja Polańskiego, Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej, informujące o powołaniu mnie na recenzenta w postępowaniu o nadanie dr. inż. Marcinowi Michalakowi stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

Dr inż. Marcin Michalak uzyskał w roku 2009 stopień doktora nauk technicznych w dyscyplinie naukowej informatyka na Wydziale Automatyki, Elektroniki i Informatyki Politechniki Śląskiej przedstawiając rozprawę „Estymacja funkcji regresji w predykcji szeregów czasowych”. Od roku 2014 pracuje jako adiunkt na Wydziale Automatyki Elektroniki i Informatyki, Politechniki Śląskiej i od roku 2016 jako starszy specjalista w Instytucie Technik Innowacyjnych EMAG (Sieć Badawcza Łukasiewicza).

Kandydat uprzednio nie ubiegał się o nadanie stopnia doktora habilitowanego.

Recenzja wniosku o nadanie stopnia doktora habilitowanego sporządzona jest zgodnie z wymaganiami określonymi w art. 219 ust. 1 pkt 2 ustawy „Prawo o szkolnictwie wyższym i nauce” (j.t. Dz.U.2020 r. poz. 85 z późn. zm.) .

Ocenę merytoryczną przeprowadzono na podstawie dokumentacji przekazanej mi przez jednostkę prowadzącą przedmiotowe postępowanie.

Ocena wskazanego przez kandydata osiągnięcia naukowego, w tym, czy stanowi ono znaczny wkład w rozwój określonej dyscypliny naukowej.

Jako osiągnięcie naukowe przedstawiono następujący cykl dziesięciu prac:

1. M. Michalak: Theoretical Backgrounds of Boolean Reasoning Based Binary n-clustering , Knowledge and Information Systems, vol. 64, str. 2171 - 2188, 2022 /IF = 2.531 (2021); 100 pkt. MEN

2. M. Michalak: Hierarchical heuristics for Boolean reasoning based binary bicluster induction, *Acta Informatica*, vol. 59(6), str. 673-685, 2022 /IF = 0.871 (2021); 70 pkt. MEN
3. M. Michalak, R. Jaksik, D. Ślęzak: Heuristic Search of Exact Biclusters in Binary Data, *International Journal of Applied Mathematics and Computer Science*, vol. 30(1), str. 161-171, 2020 /IF = 1.417; 100 pkt. MEN
4. M. Michalak, D. Ślęzak: On Boolean Representation of Continuous Data Biclustering, *Fundamenta Informaticae*, vol. 167(3), str. 193-217, 2019 /IF = 1.298; 70 pkt. MEN
5. M. Michalak, D. Ślęzak: Boolean Representation for Exact Biclustering, *Fundamenta Informaticae*, vol. 161(3), str. 275-297, 2018 /IF = 1.204; 70 pkt. MEN
6. M. Michalak: Induction of Centre-Based Biclusters in Terms of Boolean Reasoning, *Man Machine Interactions 6. ICMMI 2019*, A. Gruca et al., Eds., str. 239-248, *Advances in Intelligent Systems and Computing*, vol. 1061, Springer, Cham. /20 pkt. MEN
7. M. Michalak, M. Stawarz: 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, *Lecture Notes in Computer Science*, vol. 6922, Springer, Berlin, Heidelberg, 2011. /Web of Science 15 pkt. MEN
8. M. Michalak: 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*, 2022. /20 pkt. MEN
9. M. Michalak, R. Jaksik: 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*, 2022. /20 pkt. MEN
10. M. Michalak: 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*, 2019. /20 pkt. MEN

Z tych prac pięć zostało opublikowanych w znanych czasopismach indeksowanych w bazie *Journal Citation Reports (JCR)*, dwie prace to publikacje konferencyjne, a pozostałe trzy to rozdziały w opracowaniach lokalnych wieloautorskich.

Habilitant wskazał w dokumentacji na dominującą swoją rolę w pracach współautorskich. Nie otrzymałem w dokumentacji oświadczeń współautorów.

Sumaryczny IF publikacji wchodzących w skład omawianego cyklu prac wynosi 7,321 (ważony: 6,535), a liczba uzyskanych punktów wg MNiSW to 490 (ważona: 362). Prace były cytowane w *Web of Science* 136 razy (119 po wykluczeniu autocytowań), *index Hirscha* wynosi 7.

Te dane naukometryczne oceniam pozytywnie.

Przedstawione przez dr. inż. Marcina Michalaka osiągnięcie naukowe składające się z cyklu 10 spójnych tematycznie prac, opublikowanych w latach 2011 (1 praca) i 2018-2022 (9 prac) spełnia formalne wymogi stawiane osiągnięciom naukowym przez odpowiednie przepisy i ustawy - stanowi podstawę dla niniejszej recenzji.

Przedstawione w osiągnięciu naukowym wyniki badań można pomieścić w dziedzinie wskazywanej przez Profesora Leslie Valianta, laureata nagrody Turinga z Uniwersytetu Harvarda jako fundamentalną dla sztucznej inteligencji (por. <http://people.seas.harvard.edu/~valiant/researchinterests.htm>), w dziedzinie związanej z poszukiwaniem czy też charakteryzacją obliczeniowych bloków konstrukcyjnych niezbędnych dla zrozumienia (percepcji) postrzeganej sytuacji. Te obliczeniowe bloki budujące w obliczeniach granularnych odpowiadają granułom. Metody

odkrywania granul istotnych dla zrozumienia postrzeganej sytuacji w stopniu pozwalającym na podejmowanie właściwych decyzji należą do kluczowych w sztucznej inteligencji.

Wyniki zawarte w recenzowanym osiągnięciu naukowym związane z odkrywaniem granul, odpowiednich dla rozważanego zadania, zostały uzyskane z wykorzystaniem wnioskowania boolowskiego. George Boole, w połowie XIX wieku, zaproponował aby do rozwiązywania problemów wykorzystać funkcje boolowskie. Istota zaproponowanego podejścia sprowadza się do tego aby dla zadanego problemu P skonstruować taką funkcję boolowską f_P , której implikanty (pierwsze) dają się łatwo zinterpretować jako rozwiązania problemu P . Na przestrzeni lat ta metodologia okazała się bardzo skuteczna przy rozwiązywaniu wielu klas problemów. W szczególności, wyniki recenzowanego osiągnięcia naukowego dotyczą ważnej dla zastosowań praktycznych klasy problemów dotyczących biklasteryzacji, ogólniej n -klasteryzacji, danych i stanowią one znaczący wkład autora w rozwój metod wnioskowania boolowskiego w odniesieniu do tej klasy problemów. Należy tu podkreślić, że w związku z dużą złożonością obliczeniową problemu poszukiwania implikantów pierwszych, odpowiadających rozwiązaniom optymalnym rozważanych problemów, często stosuje się tzw. aproksymacyjne wnioskowanie boolowskie prowadzące do wyznaczania implikantów bądź też aproksymacji implikantów pierwszych za pomocą efektywnych heurystyk prowadzących do rozwiązań semi-optymalnych lub aproksymacyjnych. Przedstawione wyniki w recenzowanym osiągnięciu naukowym dotyczą tych różnych wersji wnioskowania boolowskiego w odniesieniu do szeregu problemów ze wspomnianej klasy problemów ważnej dla zastosowań. W osiągnięciu naukowym są one formułowane w postaci słabych i mocnych twierdzeń. Wśród tych problemów do najważniejszych zaliczam:

- Problem poszukiwania biklastrów dokładnych w macierzach dyskretnych.
- Problem poszukiwania biklastrów podobieństwa w macierzach ciągłych.
- Problem poszukiwania biklastrów chaosu w macierzach ciągłych.
- Problem poszukiwania biklastrów wokół ustalonej wartości centralnej i zadanej wartości rozrzutu.
- Problem poszukiwania n -klastrów w hiperkostkach binarnych.

W każdym z tych przypadków przedstawiono wersje słabe i mocne twierdzeń uzasadniających poprawność kodowania boolowskiego problemu prowadzącego do odpowiedniej funkcji boolowskiej. Ponadto przedstawiono też, pewne proste heurystyki takie jak np. heurystyka Johnsona.

Uwagi krytyczne.

1. Pewnym mankamentem jest brak głębszych badań związanych ze skalowalnością opracowanych metod. Wiadomo, że funkcje boolowskie odpowiadające rzeczywistym problemom są bardzo dużych rozmiarów i są trudne do analizy. Jako przykład mogą posłużyć zmagania wielu badaczy pracujących nad SAT solverami.
2. Bardzo duży rozmiar funkcji boolowskich odpowiadających rzeczywistym problemom wymaga opracowania metod reprezentacji tych funkcji np. w tablicach danych o stosunkowo małych rozmiarach, z których te funkcje można wygenerować. Chodzi więc o metody, które nie operują bezpośrednio na samych funkcjach ale na ich zwartych reprezentacjach. Brak jest szerszej dyskusji w osiągnięciu naukowym w związku z tym zagadnieniem.

Wymienione uwagi krytyczne nie wpływają na moją pozytywną ocenę wyników przedstawionych w osiągnięciu naukowym. Oceniam je jako spełniające wymagania stawiane przez odpowiednie przepisy odnośnie nadawania stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

Ocena pozostałych wyników naukowych

Dr inż. Marcin Michalak ma znaczący dorobek naukowy nie wchodzący w skład osiągnięcia naukowego. Po uzyskaniu stopnia doktora opublikował on 85 prac, z których 10 stanowi osiągnięcie naukowe. Wśród tych pozostałych prac jest 8 prac opublikowanych w czasopismach z listy JCR. Wśród prac Kandydata są prace opublikowane w znanych czasopismach takich jak w przypadku osiągnięcia naukowego: *Knowledge and Information Systems*, *Acta Informatica*, *Fundamenta Informaticae*, *Journal of Applied Mathematics and Computer Science*; oraz w przypadku pozostałych prac: *IEEE Access*, *Sensors*, *IET Intelligent Transport Systems*, *Maintenance and Reliability*. Oprócz tego jest on autorem 13 prac w czasopismach spoza listy JCR ale z listy ministerialnej i kilkudziesięciu prac opublikowanych w sprawozdaniach konferencyjnych oraz jako rozdziały książek.

Sumaryczny współczynnik wpływu według listy Journal Citation Reports (JCR), zgodnie z rokiem opublikowania wynosi 23,412 (ważony 12,334). Liczba cytowań publikacji według bazy Web of Science (WoS) wynosi 136 (119 bez autocytowań), liczba prac 41. Indeks Hirscha według bazy Web of Science (WoS) wynosi 7. Liczba cytowań publikacji według bazy Google Scholar wynosi 375. Indeks Hirscha według bazy Google scholar wynosi 10. Sumaryczna liczba punktów wg MEN wynosi 1501 (ważona 784).

Tematyka tych prac, oprócz zagadnień związanych z biklasteryzacją, stanowiących uzupełnienie osiągnięcia naukowego (np. o metody nie bazujące na wnioskowaniu boolowskim), obejmuje: indukcję reguł wraz z oceną ich jakości, diagnostykę urządzeń i budowę modeli diagnostycznych, analizę szeregów czasowych, analizę obiektów odstających, wizję komputerową w analizie danych i specyficzne metody analizy danych.

Do najważniejszych z tych prac zaliczam:

- (i) cykl prac dotyczący diagnostyki i modeli diagnostycznych powiązanych z rzeczywistymi projektami w których brał udział Kandydat,
- (ii) cykl prac o nieparametrycznych metodach estymacji funkcji regresji w zastosowaniu do predykcji finansowych szeregów czasowych oraz metodach estymacji funkcji regresji w przypadku szeregów czasowych z zastosowaniem autorskich funkcji jądra,
- (iii) cykl prac dotyczących analizy obserwacji odstających w rzeczywistych projektach dotyczących wykrywania potencjalnych zagrożeń w ruchu sieciowym.

Jako bardzo pozytywną cechę tych prac uznają fakt, że powstały one najczęściej przy aktywnym udziale habilitanta w różnych projektach o istotnym znaczeniu praktycznym.

Wyniki naukowe Kandydata nie objęte osiągnięciem naukowym oceniam pozytywnie.

Informacja o spełnieniu przez kandydata kryterium dotyczącego wykazywania się istotną aktywnością naukową

Jak już wspomniałem, Kandydat opublikował ponad 80 prac naukowych. Uzyskane wyniki naukowe referował na wielu konferencjach naukowych.

Na szczególne wyróżnienie zasługuje aktywność habilitanta (jako wykonawcy bądź kierownika) w wielu (25-ciu!) projektach naukowo-badawczych realizowanych zarówno z naukowcami z kraju jak i z zagranicy.

Dr inż. Marcin Michalak od roku 2016 pracuje dodatkowo jako starszy specjalista w Sieci Badawczej Łukasiewicz - Instytut Technik Innowacyjnych EMA. Aktywnie współpracował i współpracuje z naukowcami z innych ośrodków. Na przykład, dotyczy to współpracy z:

- (i) Profesorem Aguilar_Ruiz'em (Universidad Pablo de Olavide, Sewilla, Hiszpania) w zakresie biklasteryzacji i oceny klasyfikatorów (od 2019),
- (ii) Profesorem Dominikiem Ślęzakiem (Wydział Matematyki Informatyki i Mechaniki Uniwersytetu Warszawskiego) w zakresie matematycznych podstaw biklasteryzacji na bazie wnioskowania boolowskiego (2015-2020);
- (iii) pracownikami Głównego Instytutu Górnictwa: z dr inż. Karolina Nurzyńska i dr inż. Sebastian Iwaszenko w zakresie prac związanych z modelowaniem rozwoju kawerny w procesie podziemnego zgazowania węgla (2013 -2014) oraz z dr Leokadia Róg w zakresie analizy składu petrograficznego próbek węgla kamiennego.

Powyższe fakty oraz udział w wielu projektach jak i wspólne publikacje z naukowcami z różnych ośrodków w kraju jak i z zagranicy wskazują, że dr inż. Marcin Michalak spełnia z nadmiarem wymagania dotyczące współpracy z większą liczbą ośrodków naukowo-badawczych.

Dr inż. Michalak wykazał się dużą aktywnością w środowisku międzynarodowym. Wygłaszał referaty dotyczące uzyskanych wyników naukowych na ponad 30 konferencjach naukowych. Jest członkiem komitetów redakcyjnych dwóch czasopism (w *Symmetry* jako Guest Editor, w *Mining - Informatics, Automation and Electrical Engineering* jako Managing Editor). Jest często powoływany jak recenzent prac naukowych zgłaszanych do czasopism naukowych oraz na konferencje naukowe. Odbił kilka staży naukowych w ośrodkach krajowych i zagranicznych, w tym (i) kilkutygodniowy staż naukowy w Universidad Pablo de Olavide, Sewilla, Hiszpania, finansowany w ramach uzyskanego Rektorskiego Grantu Habilitacyjnego (maj/czerwiec 2022), (ii) blisko siedmioletniowy staż naukowy w Universidad Pablo de Olavide, Sewilla, Hiszpania, finansowany w ramach uzyskanego grantu na działanie naukowe Miniatura 3 (luty/marzec 2020, listopad 2021 -przerwa ze względu na sytuację pandemiczną) oraz (iii) miesięczny staż naukowy w Instytucie Technik Innowacyjnych EMAG (Katowice, Polska).

Na wyróżnienie zasługuje aktywność dr. inż. Marcina Michalaka w zakresie kształcenia młodej kadry. Był on co-promotorem w dwóch zakończonych przewodach doktorskich i jest co-promotorem w trzech przewodach doktorskich. Ponadto, jest on opiekunem naukowym projektów realizowanych przez doktorantów w ramach projektu *Utworzenie nowej specjalności w dyscyplinie informatyka Politechniki Śląskiej*.

Za działalność naukową i organizacyjną Kandydat został dwukrotnie wyróżniony nagrodą Rektora Politechniki Śląskiej. Jako członek zespołu został on wyróżniony dyplomem dla Głównego Instytutu Górnictwa w Katowicach za projekt *Przyrząd do mierzenia emisji dwutlenku węgla z hałd kopalnianych* oraz srebrnym medalem na targach Brussels Innova: The Belgian and International Trade Fair for Technological Innovation za innowację *Instrument for measuring carbon dioxide emission from mine waste dumps*. Otrzymał on również *International Fuzzy Systems Association Award for Young Scientist* na konferencji *Advances in Artificial Intelligence and Applications (FedCSIS 2015)*, wspólnie z M. Sikorą i Ł. Wróblem za artykuł *Rule Quality Measures Settings in a Sequential Covering Rule Induction Algorithm: An Empirical Approach*.

Kandydat spełnia wymagania odnośnie kryterium dotyczącego wykazania się istotną aktywnością naukową.

Informacja o osiągnięciach dydaktycznych, organizacyjnych i popularyzujących naukę kandydata do stopnia doktora habilitowanego

Dr Inż. Marcin Michalak ma znaczące doświadczenie dydaktyczne. W czasie wieloletniej pracy w Politechnice Śląskiej prowadził (w języku polskim i angielskim) zajęcia z wielu przedmiotów informatycznych, takich jak matematyka dyskretna i logika matematyczna, narzędzia i środowiska

zaawansowanej analizy danych, analiza danych i inteligencja obliczeniowe, wybrane zagadnienia matematyczne, odkrywanie wiedzy, szeregi czasowe.

Był promotorem 6 prac magisterskich i 2 inżynierskich na Wydziale Automatyki, Elektroniki i Informatyki Politechniki Śląskiej na kierunku Informatyka.

Udział Kandydata w wielu projektach naukowo-badawczych wskazuje również na poważne jego osiągnięcia organizacyjne. Ponadto, brał on udział w komitetach organizacyjnych kilku konferencji międzynarodowych.

W ramach działalności popularyzującej naukę kandydat wygłosił kilka wykładów dla uczniów II i III klasy Gimnazjum nr 2 w Mikołowie (np. o relacyjnym bazach danych (2018), budowie prostego komputera na przykładzie Maszyny W (2018), regresji liniowej metodą najmniejszych kwadratów (2017)).

Podsumowanie

Reasumując uważam, że dorobek naukowy dr. inż. Marcina Michalaka zawarty w osiągnięciu naukowym i w pozostałych publikacjach, jego aktywność naukowa, jak również jego wkład w kształcenie młodej kadry, osiągnięcia dydaktyczne i organizacyjne spełniają wymagania stawiane przez odpowiednie przepisy przy ubieganiu się o stopień naukowy doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.