

Recenzja
osiągnięcia naukowego pt.

**Sterowanie przydziałem zasobów w energooszczędnych sieciach
i rozproszonych systemach przetwarzania danych**

dr. inż. Piotra Arabasa

w związku z postępowaniem o nadanie stopnia naukowego doktora habilitowanego
w dziedzinie nauk inżynieryjno-technicznych i w dyscyplinie
informatyka techniczna i telekomunikacja

prowadzonym przez
Radę Naukową Dyscypliny Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej

1. Sylwetka Habilitanta

Dr inż. Piotr Arabas uzyskał w 1996 r. tytuł magistra inżyniera w specjalności systemy sterowania i wspomagania decyzji uzyskany w Instytucie Automatyki i Informatyki Stosowanej Politechniki Warszawskiej za pracę pt. *Rozproszona implementacja algorytmu wyznaczania optymalnych sterowań w przypadku prognoz wielowariantowych*.

W 2004 r. uzyskał stopień doktora nauk technicznych w specjalności automatyka uzyskany na Wydziale Elektroniki i Technik Informacyjnych Politechniki Warszawskiej za rozprawę pt. *Hierarchiczna struktura w systemie obrony przeciwrakietowej; mechanizmy decyzyjne i badania symulacyjne*.

W latach 2002–2004 pracował jako asystent, a od roku 2004 jako adiunkt w Instytucie Automatyki i Informatyki Stosowanej, Wydział Elektroniki i Technik Informacyjnych Politechniki Warszawskiej.

Jednocześnie w latach 2002–2004 był asystentem w Pionie Naukowym, a od 2004 adiunktem w Centrum Badań i Transferu Technologii w Naukowej i Akademickiej Sieci Komputerowej, Państwowym Instytucie Badawczym.

2. Ocena głównego osiągnięcia naukowego

Główne osiągnięcie Habilitant przedstawił w cyklu tematycznie powiązanych publikacji obejmujących jedną książkę, której jest współautorem, oraz 11 publikacji w czasopismach z listy A MNiSW i 1 publikację w materiałach konferencji z listy CORE B.

[H1] Arabas, P. (2021) Modeling and simulation of hierarchical task allocation system for energy-aware HPC clouds. *Simulation Modelling Practice and Theory*, 107:102221. DOI:10.1016/j.simpat.2020.102221. IF21=4,199, 100 pkt. MNiSW, 100%.

[H2] Arabas, P., Niewiadomska-Szynkiewicz E. (2020) Energy-efficient workload allocation in distributed HPC system. W: *The 2019 International Conference on High Performance Computing & Simulation HPCS 2019*, 470-477. DOI:10.1109/HPCS48598.2019.9188240. CORE2018 B, 70 pkt. MNiSW, 70%.

[H3] Arabas P. (2018) Energy aware data centers and networks: a survey. *Journal of Telecommunications and Information Technology*, 4/2018:26-36. DOI:10.26636/jtit.2018.129818. 12 pkt. (obecnie 40 pkt.) MNiSW, 100%

- [H4] Niewiadomska-Szynkiewicz, E., Sikora, A., Arabas, P., Kołodziej, J. (2013) Control system for reducing energy consumption in backbone computer network. *Concurrency and Computation – Practice & Experience*. 25(12):1738-1754. DOI:10.1002/cpe.2964. IF13=0,784, 25 pkt. (obecnie 100 pkt.) MNiSW, 20%.
- [H5] Niewiadomska-Szynkiewicz, E., Sikora, A., Arabas, P., Kamola, M., Mincer, M., Kołodziej, J. (2014) Dynamic power management in energy-aware computer networks and data intensive computing systems. *Future Generation Computer Systems*, 37:284-296. DOI:10.1016/j.future.2013.10.002. IF14=2,786, 40 pkt. (obecnie 140 pkt.) MNiSW, 20%.
- [H6] Arabas P., Joźwik T., Niewiadomska-Szynkiewicz, E. (2023) Router activation heuristics for energy-saving ECMP and valiant routing in data center networks, *Energies*, 16:1–20. DOI:10.3390/en16104136, IF22=3,2, 140 pkt. MNiSW, 50%.
- [H7] Karpowicz, M., Arabas, P., Niewiadomska-Szynkiewicz, E. (2018) Energy-aware multilevel control system for a network of Linux software routers: Design and implementation, *IEEE Systems Journal*, 12(1):571-582, DOI:10.1109/JSYST.2015.2489244. IF18=4,463, 35 pkt. (obecnie 140 pkt.) MNiSW, 20%.
- [H8] Jaskóła, P.; Arabas, P.; Karbowski, A. (2016) Simultaneous routing and flow rate optimization in energy-aware computer networks, *International Journal of Applied Mathematics and Computer Science*, 26(1):231-243, DOI:10.1515/amcs-2016-0016.25 pkt. (obecnie 100 pkt.) MNiSW, 35%.
- [H9] Karpowicz M., Arabas P. (2016) Server workload model identification: Monitoring and control tools for Linux, *Journal of Telecommunications and Information Technology*, (2):5-12. 12 pkt. (obecnie 40 pkt.) MNiSW, 50%.
- [H10] Arabas P., Karpowicz M. (2016) Wykorzystanie informacji z rejestrów procesora do identyfikacji modelu poboru mocy przez serwer. *Przegląd Elektrotechniczny*, (3):34-41, 14 pkt. (obecnie 70 pkt.) MNiSW, 80%.
- [H11] Karpowicz M., Arabas P., Niewiadomska-Szynkiewicz E. (2018) Design and implementation of energy-aware application-specific CPU frequency governors for the heterogeneous distributed computing systems, *Future Generation Computer Systems*, (78):302–315. DOI:10.1016/j.future.2016.05.011. IF18=5,768, 40 pkt. (obecnie 140 pkt.) MNiSW, 20%.
- [H12] Niewiadomska-Szynkiewicz E., Marks M., Arabas P., Sikora A. (2022) *Bezprzewodowe sieci czujników w internecie rzeczy. Modele - Algorytmy - Protokoły*, 292 str., PWN. 80 pkt. MNiSW, 20% .
- [H13] Arabas P., Sikora A., Szynkiewicz W. (2021) Energy-aware activity control for wireless sensing infrastructure using periodic communication and mixed-integer programming, *Energies*, 14:1–17. DOI:10.3390/en14164828. IF21=3,252 140 pkt. MNiSW, 70%.

Prócz tego Habilitant wskazał 12 publikacji ponumerowanych od U1 do U12, które nazwał uzupełniającymi, zawierającymi dodatkowe wyniki z zakresu osiągnięcia naukowego. W autoreferacie Habilitant nie zalicza ich do pozostałej aktywności naukowej. Jednocześnie nie są one włączone bezpośrednio do osiągnięcia naukowego, chociaż wspomina się o nich w opisach wyników.

Ogromny wzrost zapotrzebowania na usługi komputerowe spowodował rozbudowę dużych centrów obliczeniowych z rozbudowanymi systemami komputerowymi. Masowe użycie tych systemów prowadzi do dużego poboru energii. Tematem głównego osiągnięcia Habilitanta jest opracowanie metod zarządzania w złożonych systemach komputerowych prowadzące do zmniejszenia zużycia energii. Tematyka ta wpisuje się w aktualne trendy oszczędności energetycznej, nie wspominając o wymiarze finansowym takich rozwiązań. W odróżnieniu od głównie rozwijanych metod polegających na minimalizacji poboru energii w urządzeniach lokalnych, których przegląd przedstawiono w [H3], Habilitant rozważa kompleksową optymalizację sieci klastrów obliczeniowych. W ramach swoich badań rozważył trzy

zagadnienia: (i) wyznaczenie energooszczędnych struktur sterowania w rozproszonym systemie obliczeniowym; (ii) opracowanie metod energooszczędnego ruchu pakietów; (iii) budowa modeli poboru mocy przez procesor i system komputerowy z estymacją parametrów na podstawie przeprowadzonych eksperymentów laboratoryjnych.

Zagadnienie (i) opisano w artykułach [H1] do [H3] oraz w [U1] i [U2]. Oszczędności energii w tym przypadku można uzyskać przez odpowiedni rozdział obciążenia klastrów obliczeniowych korzystając z tego, że typowo obciążenie nie wymaga pełnego wykorzystania dostępnej mocy obliczeniowej klastrów. Pozwala to na wprowadzenie części urządzeń systemu w stan niższego poboru energii, na przykład w stan czuwania. Zadanie polega więc na takim przydziale zadań obliczeniowych, aby uzyskać możliwie największe oszczędności. W zadaniu postawionym w pracach Habilitanta rozważa się przetwarzanie danych napływających w sposób ciągły ze znaną intensywnością, przy czym napływ ten jest stały w czasie. To założenie pozwala na sformułowanie zadania przydzielenia mocy obliczeniowej do przetwarzania danych jako statycznego zadania alokacji zasobów, w którym optymalizowanym kryterium jest zużycie energii. Nowością w stosunku do wcześniejszych prac jest uwzględnienie w tym zadaniu poboru mocy przez urządzenia wchodzące w skład sieci łączącej klastry, co Habilitant opisał w artykule [H1]. Jest to zadanie liniowe mieszane ze zmiennymi całkowitoliczbowymi, którego wymiarowość w praktycznie interesujących zadaniach może być duża.

Z tego względu Habilitant zaproponował zastosowanie metody dwupoziomowego podejmowania decyzji, w którym w klastrach obliczeniowych są rozwiązywane zadania optymalizacji lokalnej, a koordynatorem jest zarządca całej sieci. Klastry podają koordynatorowi liczbę udostępnianych procesorów oraz średni pobór energii związany z ich działaniem. Na tej podstawie koordynator rozwiązuje zadanie programowania całkowitoliczbowego, o znacznie mniejszej wymiarowości niż zadania mieszanego w pełnym sformułowaniu, i przydziela klastrom zadania obliczeniowe, które w wyniku optymalizacji całkowitoliczbowej są rozdzielane w klastrze do poszczególnych procesorów. Kluczowym fragmentem użytej metody jest rozwiązanie sposobu rozdziału tych zadań skierowanych do klastrów, których nie daje się ulokować na procesorach z powodu braku ich wystarczających mocy obliczeniowych. Habilitant zaproponował heurystyczną iteracyjną metodę rozwiązania tego zagadnienia. W rezultacie opracował suboptymalny algorytm o zazwyczaj znacznie mniejszej złożoności obliczeniowej niż pełny algorytm optymalizacji mieszanej. Został on przebadany eksperymentalnie dla dwóch przykładowych topologii sieci. Względna oszczędność energii w stosunku do maksymalnego stanu energetycznego systemu wynosi dla dużych liczb zadań około 0,63 – 0,64 %, jednak dla mniejszych obciążeń zależy istotnie od topologii sieci, wzrastając dla prostych topologii i spadając dla skomplikowanych.

Podsumowując, w tej części osiągnięcia Habilitanta sformułował po raz pierwszy zadanie optymalizacji energii w sieci klastrów obliczeniowych z uwzględnieniem poboru mocy przez urządzenia sieciowe, opracował matematyczne modele optymalizacyjne, zaproponował heurystyczną metodę rozwiązania pozwalającą na uzyskanie wyniku w rozsądnym czasie nawet dla dużych zadań i przeprowadził obliczenia symulacyjne. Prezentacja materiału w podanych artykułach jest klarowna i zrozumiała. Wiodący wkład Habilitanta jest tutaj oczywisty: we wszystkich publikacjach [H1] do [H3] jest on pierwszym autorem, a w najważniejszych [H1] i [H3] jedynym autorem. Tę część osiągnięcia oceniam bardzo pozytywnie.

W zagadnieniu (ii) rozważono natomiast problem oszczędności energii przez odpowiednie trasowanie ruchu w sieci. Podobnie jak w poprzednim zadaniu, niepełne wykorzystanie sieci pozwala na takie trasowanie ruchu, aby pozwolić na wprowadzenie części urządzeń sieciowych w stan niższego poboru energii bez straty jakości przesyłu. Ta część badań była w znacznym stopniu prowadzona w ramach projektu zespołowego. Główny wkład Habilitanta w tym projekcie, podany w autoreferacie, „był związany z tworzeniem środowisk sprzętowo-

programowych do badań, opracowaniem scenariuszy testowych i prowadzeniem eksperymentów mających na celu ocenę stosowności i wydajności algorytmów opracowanych w ramach projektu”.

W pracy [H5] przedyskutowano sformułowanie zadania optymalizacyjnego dla tego zagadnienia jako wyboru aktywnych węzłów oraz jako wyboru aktywnych dróg przepływu, a także przybliżone metody ich rozwiązywania. Przedstawiono też wyniki obszernych badań eksperymentalnych tych metod. Podobnie jak w zagadnieniu (i), sformułowano zadanie optymalizacyjne obejmujące całość sieci. Nawet jeżeli jest ono przybliżone zadaniem liniowym mieszanym ze zmiennymi całkowitoliczbowymi, to jest problemem NP-pełnym. W celu doprowadzenia czasu obliczeń do praktycznie akceptowalnych wartości, zaproponowano także i tu hierarchiczny schemat sterowania ruchem w sieci. Przy czym założono, że zmiany jego natężenia są na tyle wolne, że można zastosować sterowanie odcinkami stałe, co kilkakilkanaście minut. Ten schemat przedstawiono w pracach [H4, H5, U4]. Ponieważ w takim rozwiązaniu decydującym czynnikiem jest czas obliczeń, w warstwie nadrzędnej zastosowano rozwiązanie suboptymalne, a mianowicie wybór dróg przepływu z wcześniej przygotowanych i przeliczonych tras. Natomiast w warstwie podrzędnej stany energetyczne urządzeń są wybierane przez sterowniki lokalne. Przeliczone przykłady wykazały, że powoduje to bardzo dużą redukcję zmiennych w zadaniu optymalizacyjnym, szybko rosnącą z liczbą węzłów sieci. W pracach [H5, U5, U6] przedstawiono wyniki porównania tego i innych rozwiązań, zarówno pod kątem złożoności obliczeniowej, jak i eksperymentalnie, używając środowiska badawczego opracowanego przez Habilitanta.

W pracy [H6] rozpatrzono natomiast alternatywne rozwiązanie tego zagadnienia koncentrujące się na trasowaniu ruchu. Zaproponowano w nim dwie heurystyki i w symulacjach wykazano, że dla niektórych struktur sieci oszczędność poboru mocy przy transmisji danych może wynieść nawet 20%. Do badań zaproponowanych metod Habilitant opracował także laboratoryjne środowisko eksperymentalne, w którym oszczędności energii były mierzone przez zamontowane liczniki energii. Opis środowiska i wyniki przeprowadzonych badań zostały przedstawione w pracach [H5, H7, H8, U3, U7, U8]. Wskazały one, że algorytmy wyznaczania tras doprowadziły do kilkunastoprocentowej (10-14%) redukcji poboru mocy, ale z możliwym kosztem niewielkiego zwolnienia ruchu dla dużych obciążeń (2-7%), przy czym algorytmy heurystyczne charakteryzowały się większą redukcją poboru mocy przy większym zwolnieniu ruchu.

Podsumowując, w tej części osiągnięcia opracowano i przebadano symulacyjnie i na rzeczywistym stanowisku laboratoryjnym algorytmy sterowania ruchem w sieci w celu zmniejszenia poboru energii elektrycznej potrzebnej do przesłania danych. Zastosowane metody i uzyskane wyniki były opublikowane w pięciu czasopismach [H4] – [H8] wysoko indeksowanych w najbardziej uznanych bazach czasopism naukowych. Pod tym względem jest to więc najlepiej udokumentowana część osiągnięcia. Jest to rezultat prac całego zespołu wykonującego projekt ECONET, jednak Habilitant uczestniczył w sposób istotny w pracach, szczególnie dotyczących eksperymentalnego testowania i oceny metod. Świadczy o tym udział w powstałych publikacjach, a w tym jako pierwszy autor w jednej i drugi autor w dwóch innych artykułach. Także i w tym przypadku tę część osiągnięcia należy ocenić bardzo pozytywnie.

Koncepcyjnie bliskim zagadnieniem jest opracowanie energooszczędnych algorytmów dla sieci bezprzewodowych czujników. To zagadnienie jest opisane w książce [H12] wydanej przez PWN oraz w artykule [H13], którego Habilitant jest pierwszym autorem. Występują tu podobne problemy oszczędności energetycznej dla węzłów sieci (czujniki) oraz oszczędności związane z trasowaniem. Jednak w tym przypadku oba te problemy są znacznie silniej powiązane ze sobą, gdyż zużycie energii przez czujnik jest przede wszystkim związane z odbieraniem i wysyłaniem sygnałów. Dodatkowo, czujniki są zasilane z baterii, przez co funkcjonalność sieci może się załamać przez wyczerpanie baterii w czujnikach kluczowych dla jej działania. Taka sieć

wymaga też uwzględnienia problemu kolizji występujących przy jednoczesnym nadawaniu przez kilka węzłów, co rozwiązano przez podział węzłów na rozłączne nadające grupy, w których nadające węzły zajmują różne szczeliny czasowe. Istotnym wkładem Habilitanta jest właśnie opracowanie algorytmów zarządzania zasobami energetycznymi węzłów w takiej sieci. W [H13] zaproponowano dwa nowe sformułowania zadań harmonogramowania przesyłania danych w sieci z okresowym raportowaniem (PSN – periodic sensor network) dążące do minimalizowania energii wydatkowanej: (1) łącznie przez wszystkie węzły lub (2) przez najbardziej obciążony węzeł. Prowadzą one do mieszanych zadań całkowitoliczbowych. Algorytmy przebadano symulacyjnie, co wykazało, że algorytm (1) można efektywnie stosować dla stosunkowo niedużej liczby węzłów, a algorytm (2) dla średniej liczby węzłów. Całość zagadnień związanych z minimalizacją wydatku energii w bezprzewodowych sieciach czujników została opisana w książce [H12] zawierającej 291 stron. W sumie to zagadnienie jest także istotną częścią omawianego osiągnięcia Habilitanta.

Ostatnią częścią osiągnięcia jest zagadnienie (iii), w którym opracowano modele poboru mocy przez urządzenia sieciowe, co zostało zastosowane do sterowania minimalizującego pobór mocy. Jednym z istotnych problemów związanych z tym zagadnieniem jest pomiar pobieranej mocy, potrzebny do sterowania, bez użycia drogich, specjalizowanych do tego urządzeń. W [H8] jest przedstawiony model rozpraszania energii w urządzeniach sieciowych w zależności od przepływu. Jest to model liniowy, którego współczynniki wyznaczono metodą najmniejszych kwadratów lub wprost z wykresu (w artykule brak informacji na ten temat). Model ten jednak zaniżał wyniki uzyskane z pomiarów wykonanych w czasie eksperymentu przeprowadzonego na sieci laboratoryjnej. Trudności wyznaczania poboru mocy w klastrze przedstawiono w przeglądowej pracy [H9], gdzie zaproponowano interfejsy procesora do szacowania mocy. W artykule [H10] opracowano natomiast modele poboru mocy w zależności od informacji udostępnianej w rejestrze MSR procesora Intel. Rozważono trzy przypadki: mocno obciążony serwer obliczeniowy bez komunikacji sieciowej, ruter programowy przekazujący ruch sieciowy i serwer przekodowujący strumień wideo. Przeprowadzono identyfikację parametrów modeli, a następnie sprawdzono ich dokładność na niezależnych pomiarach. O ile w rozpatrzonym zakresie wielkości dla dwóch pierwszych przypadków dostatecznie dokładne okazały się modele liniowe, to w trzecim przypadku pomiary wykazały załamanie charakterystyki dla większych mocy. Autorzy przyjęli dla tego przypadku model kwadratowy, który dostatecznie dobrze aproksymuje tę charakterystykę, chociaż wydaje się, że lepszą dokładność można by było uzyskać stosując model kawałkami liniowy. Istnieje literatura podająca metody estymacji parametrów takiego modelu, łącznie z estymacją punktów przegięcia. Wadą tego podejścia jest to, że metoda ta jest bardziej uciążliwa i trzeba najpierw wykryć, którego modelu należy użyć, gdy tymczasem w metodzie przyjętej w artykule stosuje się jeden model i regresję liniową, a charakterystyka nieliniowa może być łatwo wykryta tylko przez testowanie istotności parametrów. Trochę brakuje mi wzmianki w artykule o sposobie doboru modelu, ale przeprowadzone testy na niezależnych zmiennych przekonująco wykazują, że zaproponowane modele są wystarczająco dokładne. Z praktycznego punktu widzenia należy więc uznać postępowanie za właściwe.

Powyższe modele zostały wykorzystane w badaniach opisanych w artykule [H11], którego celem było opracowanie metody minimalizowania poboru mocy urządzenia przez sterowanie częstotliwością pracy procesora. Modele poboru mocy zostały uzupełnione modelem predykcji poboru mocy w zależności od obciążenia serwera dla wybranych częstotliwości. Do tego celu przyjęto modele stochastyczne z czasem dyskretnym typu ARMAX. Także tu nie podano informacji o badaniu istotności modelu ani o doborze struktury modelu. Potwierdzenie jakości modelu są tylko wykresy dopasowania do danych.

Oba te modele zostały użyte w zadaniu optymalizacyjnym, w którym sformułowano problem minimalizacji wskaźnika jakości będącego sumą ważoną energii potrzebnej do

wykonania programów i energii zużywanej przez serwer. Optymalizacja jest prowadzona na odcinku czasu podzielonym na dyskretne odcinki i dążącym do nieskończoności, co upraszcza zadanie przez możliwość poszukiwania rozwiązania stacjonarnego. Zmienną sterującą jest częstotliwość pracy procesora w każdym odcinku w zależności od obciążenia procesora. Numerycznie uzyskano rozwiązanie tego zadania uzależnione od wagi przyjętej we wskaźniku jakości.

Regulator oparty na tym rozwiązaniu uruchomiono w laboratoryjnej sieci komputerowej NASK-u i porównano eksperymentalnie jego działanie z trzema regulatorami częstotliwości pracy procesora istniejącymi standardowo w jądrze Linuxa: *performance*, *powersave* i *ondemad*. Pierwszy z nich dba o takie wskaźniki jakości obliczeń, jak liczba zaakceptowanych połączeń między serwerami na sekundę, liczba obsługiwanych zgłoszeń na sekundę, czy czas przetwarzania zgłoszeń. Drugi kieruje się oszczędnością energii. Trzeci, podobnie jak regulator opracowany przez autorów artykułu, stara się oszczędzać energię nie tracąc zbyt wiele na jakości obliczeń. Wykonane masowe eksperymenty wykazały, że regulator autorów uzyskał o ponad 2% lepsze wyniki od regulatora *ondemad*, zarówno jeżeli chodzi o oszczędność energii, jak i jakość obliczeń. W stosunku do pozostałych regulatorów uzyskał on oczywiście gorsze wyniki niż te, które starają się optymalizować te regulatory i lepsze od tych przez nie zaniechanych.

Podsumowując, w tej części osiągnięcia zdecydowanie wyróżnia się praca [H11] dokumentująca metodę prowadzącą do budowy regulatora częstotliwości pracy procesora, który to regulator uzyskał lepsze wyniki od standardowego regulatora w jądrze Linuxa. Jest to z pewnością ważny rezultat, do którego istotnie przyczynił się Habilitant, co wynika także z wcześniej opublikowanych badań, w których opracowano częściowe rozwiązania użyte w końcowej metodzie i których Habilitant był współautorem. W szczególności, Habilitant zdecydowanie przyczynił się do badań związanych z opracowaniem modeli szacowania poboru mocy przez elementy klastra obliczeniowego, a także był głównym wykonawcą środowiska użytego do eksperymentalnego potwierdzenia użyteczności opracowanego regulatora częstotliwości procesora.

Podsumowanie osiągnięcia. Przedstawione do oceny przez Habilitanta osiągnięcie obejmuje znaczny obszar prac związanych z metodami oszczędzania energii w sieciach informatycznych i systemach przetwarzania danych. Prace te są bardzo mocno związane z rozwiązaniami technicznymi, dosyć mocno zróżnicowanymi w poszczególnych typach układów oraz w stosowanych w nich urządzeniach. Przedstawiony dorobek w naturalny sposób dzieli się więc na części połączone nadrzędnym celem oszczędzania energii i obejmuje badania dotyczące energooszczędnego rozdziału obciążeń w klastrach obliczeniowych, energooszczędnego ruchu pakietów w sieciach przewodowych oraz energooszczędnego zarządzania przesyłaniem informacji w bezprzewodowych sieciach zasilanych bateryjnie, a także oszczędzania energii serwera przez sterowanie częstotliwością pracy procesora. Ten zróżnicowany zakres spowodował, że Habilitant przedstawił 12 artykułów i monografię jako główny dorobek i 12 artykułów prezentujących prace uszczegółowiające i uzupełniające główny dorobek. Wszystkie prace z głównej części znajdują się na liście MNiSW (jedna jest pracą pokonferencyjną z konferencji punktowanej przez MNiSW), a wiele z tych prac opublikowano w czasopiśmie o wysokich wskaźnikach bibliometrycznych, co oznacza, że wzbudzają one duże zainteresowanie i jest oznaką ich znaczenia w środowisku naukowym.

Tematyka prac jest bardzo aktualna. Systemy informatyczne pobierają już obecnie bardzo duże ilości energii, a ich gwałtowny rozwój związany z coraz to nowymi zastosowaniami spowoduje, że zagadnienie oszczędności energii nabierze jeszcze większego znaczenia.

Jedenaście z przedstawionych artykułów to prace wieloautorskie, ale w większości z nich występuje 2 lub 3 autorów. Większa liczba autorów wiąże się z pracami wykonanymi w ramach dużego projektu ECONET, co jest całkowicie naturalne. Wkład Habilitanta w całości prac jest bardzo znaczący. Jest on niewątpliwie głównym wykonawcą w pierwszej z wymienionych

powyżej części osiągnięcia, gdzie jest jedynym autorem dwóch artykułów i pierwszym autorem artykułu pokonferencyjnego. Ma też decydujący udział w wynikach uzyskanych w części trzeciej jako pomysłodawca modeli oraz wykonawca badań eksperymentalnych. Także Jego udział w badaniach w części drugiej, związanych z zarządzaniem rutingiem, przede wszystkim dotyczący opracowanie i wykonania strony eksperymentalnej oraz algorytmów dla sieci bezprzewodowej, jest bardzo istotny.

Nie ulega więc wątpliwości, że przedstawione przez Habilitanta osiągnięcie jest ważne, zarówno od strony naukowej, jak i od strony aplikacyjnej, a udział Habilitanta w uzyskanych wynikach całkowicie uzasadnia stwierdzenie, że osiągnięcie jest znacznym wkładem Habilitanta w rozwój dyscypliny *informatyka techniczna i telekomunikacja*.

3. Ocena pozostałych osiągnięć naukowych

Ustawa z dnia 20 lipca 2018 r. (z późniejszymi poprawkami) wymaga, aby Habilitant „posiadał w dorobku osiągnięcia naukowe”, co według interpretacji Rady Doskonałości Naukowej oznacza, że powinny to być przynajmniej dwa osiągnięcia. Habilitant w zgłoszeniu podał trzy grupy prac poza głównym osiągnięciem, które można potraktować jako wspólne osiągnięcie powiązane tematycznie.

Pierwsza z nich dotyczy sieci społecznych, ich modelowania oraz zastosowań. Grupa związanych z tą tematyką publikacji składa się z czterech pozycji:

[P1] Kamola M., Arabas P. (2018) *Sieci społeczne i technologiczne. Jak zrozumieć, jak wykorzystać*. PWN, Warszawa, 254 str. - 80 pkt. MNiSW.

[P2] Kamola M., Arabas P. (2020) Improving time-series demand modeling in hospitality business by analytics of public event datasets. *IEEE Access*, vol. 8, 53666-53677, 100 pkt. MNiSW, IF=3,367.

[P4] Gruszczyński W., Arabas P. (2016) Application of social network inferred data to churn modeling in telecoms. *Journal of Telecommunications and Information Technology*, nr 2, 77-86, 12 (obecnie 40) pkt. MNiSW.

[P5] Kamola M., Arabas P. (2014) Network resilience analysis: Review of concepts and a country level case study. *Computer Science*, vol. 15, nr 3, 311-327, 12 (obecnie 40) pkt. MNiSW.

Najważniejszą pozycją jest tutaj książka [P1] traktująca szeroko zagadnienie sieci społecznych i technologicznych. W szczególności jest ona nastawiona na przeszukiwanie internetu i wykorzystywanie znalezionych tam informacji do modelowania. Informacje te charakteryzują się na ogół wielką liczbą niejednorodnych, niepełnych, a czasami także nieprecyzyjnych danych. Książka jest związana z wykładem prowadzonym na Politechnice Warszawskiej, jednak jest nowością pod względem objętego materiału nie tylko w Polsce, ale także w sporej mierze w literaturze angielskojęzycznej. Z tego względu uważam, że może być włączona do wykazu osiągnięć Habilitanta. Książka zawiera wprowadzenie teoretyczne dotyczące grafów, sieci i sieci złożonych, modeli sieci i semantycznych atrybutów węzłów. Następnie są omówione zagadnienia przeszukiwania i wnioskowania w sieciach, w tym dokumentów w językach naturalnych, a także przetwarzania zasobów sieciowych. Na koniec przedstawiono przykładowe zastosowania omówionych metod. Trzy z nich zostały szerzej opisane w artykułach [P2, P4, P5], gdzie przedstawiono nowe rozwiązania wybranych praktycznych problemów korzystając z zasobów sieciowych. I tak, w [P2] zaproponowano model szacowania popytu na usługi hotelarskie opracowany w ramach projektu POIR dla firmy YieldPlanet. Model składa się ze składowej estymowanej na podstawie historycznych danych o typowym zamawianiu miejsc i rzeczywistej zajętości pokoi oraz drugiej składowej, zależnej od dodatkowych informacji o wydarzeniach, takich jak targi, festiwale, czy inne imprezy. Do wyznaczenia tej drugiej składowej użyto informacji tekstowych ze stron internetowych o bardzo zróżnicowanej postaci. Ten fragment jest główną nowością

opracowanego systemu. Model sprawdzono dla sieci modeli w Krakowie i we Wrocławiu. Uzyskano o 4% lepszą predykcję w dniach wydarzeń. W [P4] opracowano natomiast model przewidywania odejść klientów w firmach telekomunikacyjnych. Zastosowano w nim segmentację klientów utworzoną na podstawie sieci ich połączeń. Ta segmentacja została użyta w hybrydowym modelu regresji, składającym się z dziewięciu modeli logistycznych. Ten model uzyskał lepszy wynik niż model uczony na całym zbiorze użytkowników. Z kolei w [P5] opisano nową metodę rekonstrukcji grafu połączeń internetowych, na poziomie krajowym, między instytucjami, zwanymi systemami autonomicznymi, pod kątem niezawodności sieci. Do tego celu użyto danych dostępnych na stronach internetowych odpowiednich organizacji. Co typowe dla tego typu danych, są one niekompletne i przedstawiają inne przekroje danych, skąd konieczne było ich ujednolicenie. Graf połączeń posłużył do analizy odporności sieci na awarie.

Książka jest cenną pozycją w polskiej literaturze, w której istnieją właściwie książki związane bardziej z teoretycznymi podstawami grafów i sieci. Tematyka przeszukiwania i korzystania z zasobów sieciowych jest bardzo aktualna i przyszłościowa. Cenne są także przykłady ukazujące, jak można praktycznie stosować uzyskaną z internetu wiedzę do rozwiązywania rzeczywistych problemów. Warto podkreślić też elementy nowości w tych przykładowych rozwiązaniach. Ten fragment prac oceniam bardzo dobrze.

Druga grupa jest związana z bezpieczeństwem sieciowym. Habilitant odwołuje się tutaj do dwóch artykułów:

[P3] Arabas P., Karpowicz M. (2016) Częstość występowania wybranych triad w sieci połączeń między systemami autonomicznymi jako wskaźnik niektórych typów anomalii ruchu. *Przegląd Telekomunikacyjny – Wiadomości Telekomunikacyjne*, nr 8-9, 1179-1184, 9 (obecnie 20) pkt. MNiSW.

[P9] Skłodowski J., Arabas P. (2023) Wykorzystanie drzew sufiksowych do efektywnej prezentacji podobieństw sesji z systemu pułapek honeypot. *Cybersecurity and Law*, nr 1 (9), 298-315, 70 pkt. MNiSW.

Są one związane z projektami opracowania i wdrożenia systemów detekcji i analizy ataków sieciowych prowadzonych w NASK-u: systemu ARAKIS-Enterprise, który jest przeznaczony dla instytucji zarządzających infrastrukturą krytyczną, niejawnego systemu ARAKIS-GOV, który jest rozwinięciem poprzedniego systemu dla instytucji państwowych objętych Krajowym Systemem Cyberbezpieczeństwa i systemu FLDX służącego do wykrywania i tłumienia ataków DDoS. Habilitant opracował dla systemu ARAKIS architekturę klastra obliczeniowego, architekturę sieciową i architekturę podsystemu analizy złośliwego oprogramowania oraz przeprowadził analizy, do czego zbudował dedykowaną instalację laboratoryjną. Kierował zespołem prowadzącym analizy. Podobne prace wykonał dla systemu FLDX. Kierował też zadaniem wykorzystania metod sztucznej inteligencji do analizy danych w projekcie NCBiR Cybermine, w którym opracowano i wdrożono system monitorowania sieci informatycznych i przemysłowych w jednej z kopalń Jastrzębskiej Spółki Węglowej. Szczegóły tych rozwiązań nie mogą być oczywiście publikowane. Artykuły [P3] i [P9] prezentują tylko pewne fragmenty prowadzonych badań. W pierwszym z nich zaadaptowano koncepcje wprowadzone w sieciach społecznościowych i przedstawiono pomysł wykrywania pewnego typu triad (połączeń między trzema węzłami sieci) jako oznaki ataku DDoS. W drugim opisano prototyp systemu analizującego dane z sieci pułapek (honeypots), w szczególności danych z sesji ssh i wyszukiwania w nich wzorców za pomocą drzew sufiksowych. W obu przypadkach podano przykłady zastosowań oparte na rzeczywistych danych.

Ten fragment działalności oceniałbym przede wszystkim pod kątem systemów cyberbezpieczeństwa opracowanych i wdrożonych w najważniejszych instytucjach państwowych. Jest to z pewnością znaczne osiągnięcie, w którym Habilitant miał znaczny udział. A w zaprezentowanych rozwiązaniach widać elementy nowości.

Ostatni fragment jest poświęcony badaniom w zakresie dynamicznej alokacji połączeń. Składa się on z trzech prac.

[P6] Kamola M., Arabas P. (2011) Dynamically established transmission paths in the future Internet – proposal of a framework. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, Vol. 59, nr 3, 357-366 (obecnie 100) pkt. MNiSW.

[P7] Tarasiuk H., Śliwiński J., Arabas P., Jaskóła P., Góralski W. (2011) Performance evaluation of signaling in the IP QoS system. *Journal of Telecommunications and Information Technology*, nr 3/2011, 12-20, 7 (obecnie 40) pkt. MNiSW.

[P8] Kamola M., Arabas P. (2009) Wykorzystanie technologii Vecta Star do przekazu audiowizualnego wysokiej rozdzielczości. *Przegląd Telekomunikacyjny – Wiadomości Telekomunikacyjne*, nr 8-9, 1508-1513, 4 (obecnie 20) pkt. MNiSW.

Artykuł [P6] zawiera propozycję globalnego modelu rynku kontraktów usług sieciowych. Rozważono w nim proces zestawiania drogi transmisji danych w sieci IP z zachowaniem wymaganej jakości i w hierarchicznym układzie dostawców usług typowym dla Internetu. Zarysowano także metodę obliczania kosztów i zysków oraz negocjacji na rynku oligopolistycznym, z niewielką liczbą dostawców.

W artykule [P7] opisano symulator połączeń użytkowników w sieci IP z wymaganiami jakościowymi. Przedstawiono koncepcję działania symulatora oraz wyniki badań jego działania przeprowadzone na opracowanym zestawie laboratoryjnym, dla dwóch typów rozkładów prawdopodobieństw generowania odstępów czasu między połączeniami i czasów połączeń. Również artykuł [P8] ma charakter badań laboratoryjnych. Przebadano w nim możliwość użycia technologii VectaStar do przesyłania transmisji wideo za pomocą sieci radiowej. Utworzono stanowisko badawcze łączące dwie lokalizacje sieci NASK: w Warszawie i Pruszkowie. Wyniki wykazały, że takie łącze sprawdza się dobrze dla strumieni HD-ready, natomiast przekaz full HD powoduje zbytnią zajętość łącza i uniemożliwia zachowanie jakości przesyłu dla innych użytkowników.

Oprócz pierwszego artykułu, przedstawione prace mają charakter badań laboratoryjnych, związanych z projektami MNiSW. Mają one wydźwięk mocno techniczny. Należy je jednak rozpatrywać łącznie z dwoma wcześniej opisanymi grupami prac, jako ich uzupełnienie. W sumie, te trzy powyższe grupy prac tworzą z pewnością drugie osiągnięcie wymagane w ustawie.

Habilitant opublikował w sumie 62 prace, w tym 2 monografie naukowe, 9 rozdziałów w monografiach naukowych, 9 artykułów indeksowanych w bazie JCR i 27 recenzowanych artykułów spoza tej bazy oraz 15 recenzowanych artykułów w materiałach konferencyjnych. Łączny współczynnik IF prac wchodzących w skład głównego osiągnięcia wynosi 23,668, a wszystkich prac 27,035. Podana we wniosku liczba cytowań według Web of Science wynosi 180, według Scopus 223, a według Google Scholar 287. Indeks Hirscha to 7 według Web of Sciences, 8 według Scopus i 10 według Google Scholar.

Warto też wspomnieć o licznych projektach (jedenastu), w których Habilitant brał udział. Były to projekty finansowane przez 5. i 7. Program Ramowy Unii Europejskiej, Europejski Fundusz Rozwoju Regionalnego, programy COST, POIR, NCBiR, NCN i MNiSW, a także sześciu wdrożeniach.

4. Informacja o aktywności dydaktycznej i organizacyjnej Habilitanta

Ustawa nie wymaga oceny tych aktywności, jednak zwyczajowo wspomina się o nich w ocenie Habilitanta.

Habilitant prowadził szeroką działalność dydaktyczną. Prowadził następujące wykłady i laboratoria:

Wirtualne środowiska obliczeniowe – wykład dla studentów II stopnia, który opracował i uruchomił w 2022 r.;

Techniki analizy sieci społecznych – wykład dla studentów II stopnia, którego jest kierownikiem, od 2015 r.;

Sieci i sterowanie systemami – wykład dla studentów II stopnia prowadzony w latach 2014 – 2022;

Programowanie obiektowe – wykład dla studentów I stopnia w latach 2008 – 2020;

Sterowanie sieciami komputerowymi – wykład dla studentów II stopnia w latach 2005-2007;

Sieci komputerowe – wykład i laboratoria dla Wydziału Inżynierii Produkcji Szkoły Głównej Gospodarstwa Wiejskiego w roku akademickim 2005-2006. Kierownik przedmiotu i autor projektu laboratorium;

Sieci komputerowe – laboratorium dla studentów I stopnia prowadzone w latach 2003 – 2022, Computer Networks – laboratorium dla studentów I stopnia Wydziału Matematyki i Nauk Informacyjnych prowadzone w latach 2003 – 2016;

Computer Networks – laboratorium dla studiów w języku angielskim prowadzone w latach 2003 - 2016.

Prowadził prace organizacyjne jako:

Kierownik Zespołu Inżynierii Systemów Sieciowych NASK – w latach 2018-2022;

Kierownik Laboratorium Zespołu Inżynierii Systemów Informatycznych NASK w latach 2016-2018;

Sekretarz seminarium w Zespole Złożonych Systemów;

Kierownik Zadania w projekcie NCBiR „Centrum monitorowania instalacji przemysłowych w podziemnych zakładach górniczych i wykrywania cyberzagrożeń” POIR.01.01.01-00-0180/22 – w roku 2023;

Członek komitetu naukowego konferencji: High Performance Modeling and Simulation w latach 2015-2017, 22nd International Teletraffic Congress, Specialist Seminar on Energy Efficient and Green Networking, Technologies and Materials for Renewable Energy, Environment and Sustainability.

Wykonał kilkadziesiąt recenzji dla czasopism, m.in. *IEEE Journal on Selected Areas in Communications*, *IEEE Transactions on Services Computing*, *Sensors* oraz konferencji naukowych, a także kilkadziesiąt recenzji prac dyplomowych na Wydziale Elektroniki Technik Informacyjnych PW.

Uzyskał ważne nagrody:

Nagroda Lider Bezpieczeństwa Państwa 2022 za system Arakis Enterprise;

Nagroda zespołowa I stopnia Rektora PW za osiągnięcia naukowe w roku 2022;

Nagroda zespołowa III stopnia Rektora PW za osiągnięcia dydaktyczne w r. akad. 2014/2015;

Nagroda zespołowa I stopnia Rektora PW za osiągnięcia naukowe w latach 2012-2013;

Nagroda indywidualna II stopnia Rektora PW za osiągnięcia naukowe w roku 2004;

Nagroda zespołowa III stopnia Rektora PW za osiągnięcia naukowe – prace w projekcie 5. Programu Ramowego UE Quality of Service and Pricing Differentiation for IP Services, 2003;

Nagroda „Best Paper Award” za artykuł: Niewiadomska-Szynkiewicz, E., Sikora, A., Arabas, P., Malinowski K., Energy-saving management in computer networks, *Australian Journal of Electrical and Electronics Engineering*, vol. 12, nr 3, 242-252, 2015, DOI:10.1080/1448837X.2015.1093002.

Prowadził także działalność popularyzatorską.

5. Ocena spełnienia wymagań ustawowych

Habilitant:

1) posiada stopień doktora;

2) posiada w dorobku osiągnięcia naukowe, stanowiące znaczny wkład w rozwój określonej dyscypliny, w tym 2 cykle powiązanych tematycznie artykułów naukowych opublikowanych

w czasopiśmie naukowych lub w recenzowanych materiałach z konferencji międzynarodowych, które w roku opublikowania artykułu w ostatecznej formie były ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 ust. 2 pkt 2 lit. b;

3) wykazuje się istotną aktywnością naukową realizowaną w więcej niż jednej uczelni lub instytucji naukowej (Instytut Automatyki i Informatyki Stosowanej, Wydział Elektroniki i Technik Informacyjnych Politechniki oraz Centrum Badań i Transferu Technologii, Naukowa i Akademicka Sieć Komputerowa, Państwowy Instytut Badawczy);

4) tematyka osiągnięć i pozostałe publikacje bardzo dobrze mieszczą się w dyscyplinie *informatyka techniczna i telekomunikacja*.

Tak więc Habilitant spełnia wszystkie wymagania ustawowe.

6. Wniosek końcowy

Uważam, że Habilitant spełnia wymagania ustawowe i zwyczajowe stawiane w postępowaniu habilitacyjnym. W szczególności, Habilitant przeprowadził badania istotne dla rozwoju informatyki i opublikował prace w uznanych czasopiśmie. Pracował w projektach aplikacyjnych i wdrożeniowych, co jest istotne dla osób pracujących w dziedzinie nauk inżyniersko-technicznych. Wniosuję do Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej o dopuszczenie Habilitanta do dalszych etapów postępowania habilitacyjnego.

Zbigniew Nakorski

