

Autoreferat przedstawiający opis dorobku i osiągnięć naukowych

Dane osobowe

dr inż. Jarosław Jankowski

Posiadane dyplomy i stopnie naukowe

2013 - Uniwersytet Szczeciński, dwusemestralne studia podyplomowe z zakresu zarządzania projektem badawczym i komercjalizacji wyników badań.

2003 - Politechnika Szczecińska, Wydział Informatyki, doktor nauk technicznych w dyscyplinie informatyka, specjalność: systemy wspomagania decyzji, Temat rozprawy doktorskiej: *Wielokryterialna optymalizacja kampanii reklamowych w Internecie*.

1997 - Politechnika Szczecińska, Wydział Informatyki, magister inżynier informatyki, specjalność: systemy wspomagania decyzji.

Informacje o dotychczasowym zatrudnieniu

2009 - obecnie, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Wydział Informatyki, Katedra Inżynierii Systemów Informacyjnych, stanowisko: adiunkt, od 2013 r. kierownik Zakładu Systemów Informatycznych przekształconego w 2016 r. w Zakład Analityki Systemów Internetowych i Przetwarzania Danych, od 2016 r. kierownik Katedry Inżynierii Systemów Informacyjnych.

2004 - 2009, Politechnika Szczecińska, Wydział Informatyki, Instytut Systemów Informatycznych, stanowisko: adiunkt.

1997 - 2004, Politechnika Szczecińska, Wydział Informatyki, Instytut Systemów Informatycznych, stanowisko: asystent.

Profile internetowe

Strona internetowa: jjankowski.zut.edu.pl

ResearchGate: https://www.researchgate.net/profile/Jaroslaw_Jankowski

GoogleScholar: <https://scholar.google.pl/citations?user=xZJ38c4AAAAJ>

DBLP: <http://dblp.uni-trier.de/pers/hd/j/Jankowski:Jaroslaw>

Osiągnięcie naukowe

Jako osiągnięcie naukowe w dyscyplinie informatyka wskazuję cykl trzynastu powiązanych tematycznie publikacji p.t. **Modelowanie przestrzeni wewnętrznej internetowych systemów społecznościowych**. Cykl ten obejmuje osiem artykułów opublikowanych w czasopismach naukowych z listy Journal Citation Reports oraz pięć artykułów wydanych w recenzowanych materiałach konferencyjnych wydawców takich jak: Springer oraz IEEE indeksowanych w bazie Web of Science. W skład cyklu publikacji wchodzi następujące prace:

- A1 Jankowski, J., Bródka, P., Kazienko, P., Szymanski, B., Michalski, R., Kajdanowicz, T.: Balancing Speed and Coverage by Sequential Seeding in Complex Networks**, Scientific Reports, Nature Publishing Group, 2017, VOL DOI: 10.1038/s41598-017-00937-8
IF = 5.228, 40 pkt. MNiSW, udział 50%
Wkład autorski: wiodący udział w opracowaniu koncepcji algorytmów sekwencyjnego inicjowania procesów rozprzestrzeniania informacji w sieciach złożonych i rozszerzeń związanych z ich wznawianiem oraz buforowaniem, zaplanowanie eksperymentów, opracowanie środowiska symulacyjnego, przeprowadzenie symulacji, implementacja referencyjnych algorytmów i porównanie ich efektywności, znaczący udział w analizach danych oraz opracowaniu manuskryptu.
- A2 Jankowski, J.: Dynamic Rankings for Seed Selection in Complex Networks**, Entropy 2017, 19(4), 170; doi:10.3390/e19040170
IF = 1.743, 30 pkt. MNiSW, udział 100%
- A3 Jankowski, J., Michalski, R.: Increasing Coverage of Information Spreading in Social Networks with Supporting Seeding**, The Second International Conference on Data Mining and Big Data, IEEE #41362, 07.2017, Springer, LNCS, 2017, w druku
Indeksacja w WebOfScience, 15 pkt. MNiSW, udział 90%
Wkład autorski: koncepcja badań, opracowanie algorytmów oddziaływania wspomagającego, zaplanowanie eksperymentów, opracowanie środowiska symulacyjnego, przeprowadzenie symulacji, wiodący udział w analizach danych i opracowaniu manuskryptu.
- A4 Jankowski, J.: Mixture Seeding for Sustainable Information Spreading in Complex Networks**, Asian Conference on Intelligent Information and Database Systems, ACIIDS'2017, Springer, LNCS, vol. 10191, pp. 191-201, 2017
Seria LNCS indeksowana w WebOfScience, 15 pkt. MNiSW, udział 100%
- A5 Jankowski, J., Michalski, R., Kazienko, P. : Compensatory Seeding in Networks with Varying Availability of Nodes**, Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2013, ACM, pp. 1242-1249, 2013
Indeksacja w WebOfScience, 15 pkt. MNiSW, udział 80%
Wkład autorski: koncepcja kompensacyjnego algorytmu wyboru węzłów początkowych, zaplanowanie eksperymentów, opracowanie i implementacja algorytmów, wiodący udział w analizach danych i opracowaniu manuskryptu w części eksperymentalnej

- A6** Różewski, P., Jankowski, J., Bródka, P., Michalski, R.: **Knowledge workers' collaborative learning behavior modeling in an organizational social network**, Computers in Human Behavior, vol. 51, pp. 1248-1260, 2015
IF = 2.88, 40 pkt. MNiSW, udział 25%
Wkład autorski: udział w opracowaniu koncepcji badań i założeń modeli teoretycznych, znaczący udział w implementacji modeli w środowisku symulacyjnym, znaczący udział w realizacji badań symulacyjnych, udział w analizach analiz danych i opracowaniu manuskryptu.
- A7** Różewski, P., Jankowski, J.: **Model of Multilayer Knowledge Diffusion for Competence Development in an Organization**, Mathematical Problems in Engineering. Article ID 529256, 2015
IF = 0.644, 25 pkt. MNiSW, udział 35%
Wkład autorski: udział w opracowaniu założeń modeli teoretycznych z udziałem sieci wielowarstwowych, udział w opracowaniu planu badań i określeniu przestrzeni eksperymentalnej, znaczący udział w implementacji modeli i w przeprowadzeniu badań symulacyjnych, udział w analizie danych i opracowaniu manuskryptu.
- A8** Jankowski, J., Michalski, R., Bródka, P., Kazienko, P., Utz, S.: **Knowledge acquisition from social platforms based on network distributions fitting**, Computers in Human Behavior, vol. 51, pp. 685-693, 2015
IF = 2.88, 40 pkt. MNiSW, udział 40%
Wkład autorski: koncepcja próbkowania sieci zorientowana na dopasowanie charakterystyk próbki do rozkładów parametrów sieciowych, implementacja testowego modułu w środowisku rzeczywistym, znaczący udział w opracowaniu algorytmów próbkowania, analizach danych i opracowaniu manuskryptu.
- A9** Jankowski, J., Michalski, R., & Kazienko, P.: **The Multidimensional Study of Viral Campaigns as Branching Processes**, International Conference on Social Informatics, SocInfo'2012, LNCS, vol. 7710, Springer Berlin Heidelberg, pp. 462-474, 2012
Indeksacja w WebOfScience, 15 pkt. MNiSW, udział 80%
Wkład autorski: koncepcja wykorzystania podejścia generacyjnego w analizie i predykcji zasięgu procesów rozprzestrzeniania informacji, analizy danych z kampanii rzeczywistych, opracowanie i implementacja algorytmu predykcyjnego, przeprowadzeniem badań symulacyjnych, znaczący udział w analizach danych symulacyjnych i opracowaniu manuskryptu.
- A10** Jankowski, J., Bródka, P., Hamari, J.: **A Picture is Worth a Thousand Words: an Empirical Study on the Influence of Content Visibility on Diffusion Processes within a Virtual World**, Behaviour & Information Technology, vol. 35(11), pp. 926-945, 2016
IF = 1.211, 25 pkt. MNiSW, udział 50%
Wkład autorski: koncepcja badań eksperymentalnych, implementacja mechanizmów transmisji produktów cyfrowych i realizacja eksperymentów w środowisku rzeczywistym, znaczący udział w analizach statystycznych oraz opracowaniu manuskryptu.
- A11** Utz, S., Jankowski, J.: **Making "friends" in a virtual world: The role of preferential attachment, homophily, and status**, Social Science Computer Review, vol. 34(5), pp. 546-566, 2016

IF = 1.364, 30 pkt. MNiSW, udział 25%

Wkład autorski: koncepcja realizacji badań, implementacja systemów pomiarowych w środowisku rzeczywistym, przetwarzanie danych i wstępne analizy, opracowanie części manuskryptu związanej z opisem eksperymentów oraz udział w końcowym opracowaniu całości manuskryptu.

- A12 Jankowski, J., Kolomvatsos, K., Kazienko, P., Wątróbski, J. : Fuzzy Modeling of User Behaviors and Virtual Goods Purchases in Social Networking Platforms, Journal of Universal Computer Science, vol. 22(3), pp. 416-437, 2016**

IF = 0.546, 15 pkt. MNiSW, udział 70%

Wkład autorski: koncepcja badań i systemu, akwizycja danych z systemu rzeczywistego i ich wstępne przetwarzanie, wiodący udział w realizacji badań z wykorzystaniem rozmytego modelu wnioskowania, analizie i interpretacji wyników oraz opracowaniu manuskryptu.

- A13 Jankowski, J.: Analysis of Multiplayer Platform Users Activity Based on the Virtual and Real Time Dimension, International Conference on Social Informatics, SocInfo'2011, Springer, LNCS, vol. 6984, pp. 312-315, 2011**

Indeksacja w WebOfScience, 15 pkt. MNiSW, udział 100%

W Tabeli 1. Przedstawiono sumarycznie wskaźniki bibliometryczne publikacji powiązanych z osiągnięciem naukowym.

Tabela 1. Wskaźniki bibliometryczne publikacji wchodzących w skład osiągnięcia naukowego

Nazwa wskaźnika	Wartość
Sumaryczny wskaźnik Impact Factor	16.50
Suma punktów MNiSW	320
Suma punktów MNiSW według wkładu autorskiego	184.25

Działalność naukowa

Moje zainteresowania badawcze od rozpoczęcia pracy naukowej na Politechnice Szczecińskiej w 1997 r. (przekształconej w 2009 r. w Zachodniopomorski Uniwersytet Technologiczny) powiązane były z systemami wspomagania decyzji, interakcją człowiek-komputer oraz systemami internetowymi. Prace badawcze nawiązywały do systemów internetowych poczynając od ich pierwszych zastosowań, poprzez nurt Web 2.0, aż po rozwój platform społecznościowych do chwili obecnej. Realizowane badania wpisują się w obszary informatyki uwzględnione w klasyfikacji ACM jako informatyka zorientowana na czynnik ludzki (ang. *human-centered computing*) wśród trzynastu głównych nurtów informatyki, szczególnie w obszarze interakcji człowiek-komputer, mediów społecznościowych oraz sieci społecznych. Realizowane przez mnie badania były związane z optymalizacją systemów internetowych, zastosowaniami metod wspomagania decyzji

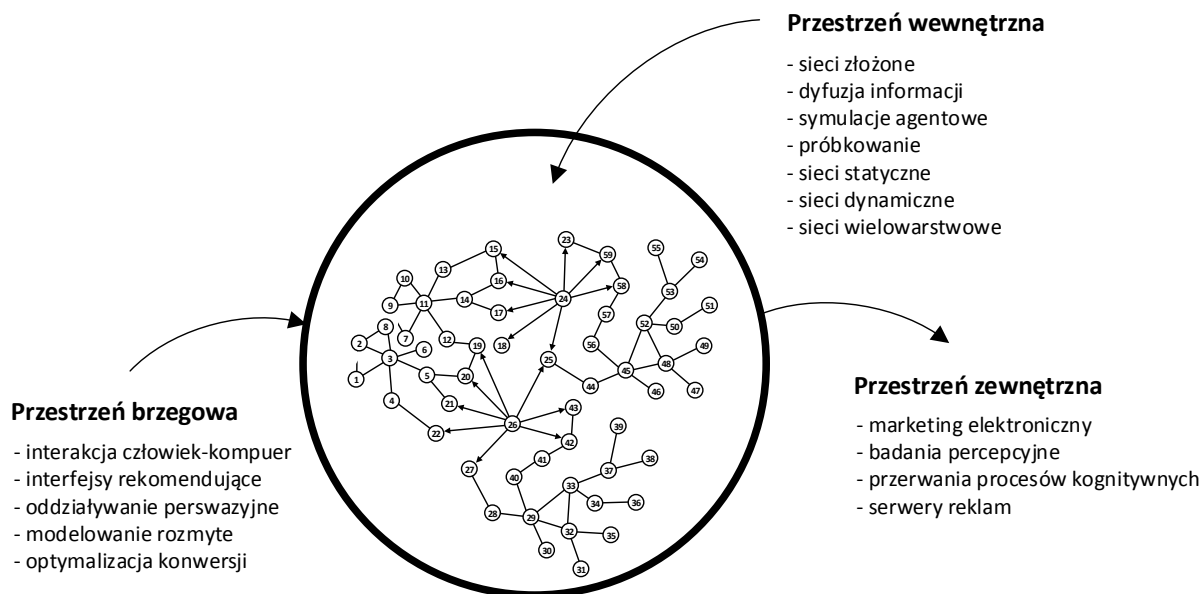
w projektowaniu systemów, a także z modelowaniem systemów społecznościowych i procesów dyfuzji informacji. Początkowe badania, również te w języku angielskim, były publikowane w wydawnictwach krajowych, a następnie wraz z tendencją umiędzynarodowienia badań nastąpiło ukierunkowanie na wydawnictwa i konferencje międzynarodowe. W badaniach wprowadzono m.in. rozwiązania zorientowane na modelowanie i eksploatację systemów społecznościowych oraz analizy zachodzących w nich zjawisk, takich jak propagacja informacji, czy treści cyfrowych. Podstawę formalną stanowiły agentowe modele symulacyjne, metody wspomagania decyzji, modelowanie rozmyte, systemy wnioskowania oraz wnioskowanie statystyczne. Poza aspektem badawczym prace mają charakter stosowany i były realizowane również we współpracy z przedsiębiorstwami z sektora internetowego w ramach wspólnych projektów oraz staży dla naukowców.

Omówienie celu naukowego wyżej wymienionych prac i osiągniętych wyników wraz omówieniem ich wykorzystania

Internetowe systemy społecznościowe łączą aspekty techniczne, społeczne jak i ekonomiczne oraz odgrywają coraz większą rolę w różnych obszarach funkcjonowania gospodarki i społeczeństw [4]. Potwierdza to dynamiczny rozwój serwisów takich jak Facebook, Twitter i wielu innych, z których korzystają setki milionów użytkowników na całym świecie. Rozwojowi serwisów tego typu towarzyszy gwałtowny przyrost danych, zarówno ustrukturalizowanych w relacyjnych oraz obiektowych bazach danych, jak i tekstowych oraz multimedialnych, które są powiązane z różnymi formami aktywności użytkowników, wykorzystaniem gier, treści multimedialnych, systemów transakcyjnych, wirtualnych produktów i usług cyfrowych [6]. Przykładowo w systemie Facebook każdego dnia użytkownicy wprowadzają ponad 250 milionów fotografii [7]. W efekcie powstają złożone struktury danych w postaci wzajemnych powiązań i relacji reprezentowanych przez sieci społeczne. Na podstawie aktywności w obrębie systemu można wnioskować o zaangażowaniu użytkowników, które łączy się z przeglądaniem profili znajomych, komentowaniem zdjęć i wpisów oraz aktualizacjami statusu [8].

Powstanie i rozwój systemów społecznościowych zwiększyły w informatyce udział czynnika ludzkiego (ang. *human centered computing*) [9], a dane związane z wzorcami zachowań użytkowników rejestrowane w obrębie systemów internetowych wykorzystywane są w ich analizach, projektowaniu i usprawnianiu. Stanowią one istotny element w cyklu życia oprogramowania. To właśnie wysoka liczba użytkowników decyduje nie tylko o wycenach przedsięwzięć internetowych, lecz także o ich potencjale gospodarczym oraz możliwościach rozwoju [10].

Z tego względu istotne są badania, które mają na celu zapewnienie większej użyteczności systemów, poprawę szeroko pojętych doświadczeń użytkownika (ang. *user experience*) oraz zwiększanie efektywności systemów, zarówno pod względem technicznym jak i ekonomicznym, przy jednoczesnym zachowaniu użyteczności dla odbiorców. Podczas ich rozpatrywania można uwzględnić podział na przestrzeń wewnętrzną, brzegową i zewnętrzną. Schematyczną strukturę internetowego systemu społecznościowego z podziałem na trzy wymienione przestrzenie wraz z obszarami badań z nimi powiązanymi przedstawiono na Rys. 1.



Rys. 1. Struktura internetowego systemu społecznościowego z podziałem na przestrzeń wewnętrzną, brzegową i zewnętrzną

Celem cyklu powiązanych tematycznie publikacji było opracowanie rozwiązań zorientowanych na modelowanie przestrzeni wewnętrznej internetowych systemów społecznościowych oraz zachodzących w nich zjawisk. Jej podstawę stanowią struktury sieci społecznych wykorzystywane jako medium propagacji informacji, produktów wirtualnych oraz innych treści. W ramach przestrzeni wewnętrznej zintegrowane są też główne mechanizmy platform społecznościowych powiązane z komunikacją, czy nawiązywaniem relacji społecznych. Złożoność obserwowanych zjawisk uzasadnia prowadzenie badań zorientowanych na pozyskiwanie wiedzy o ich specyfice oraz wprowadzanie nowych rozwiązań analitycznych. Modelowanie przestrzeni wewnętrznej internetowych systemów społecznościowych przy udziale systemów i modeli agentowych ma na celu odzwierciedlenie zachodzących w nich zjawisk, na przykład w celu zwiększania dynamiki i zasięgu procesów dyfuzji informacji, zwiększania ich efektywności, lepszego poznania ich charakterystyk oraz weryfikacji złożoności obliczeniowej algorytmów przed zastosowaniem w systemach rzeczywistych.

Procesy rozprzestrzeniania się informacji są podstawą marketingu wirusowego [3], zmian społecznych i politycznych [5], dyfuzji innowacji [11] oraz innych inicjatyw z podłożem emocjonalnym [12]. Treści przekazywane w systemach elektronicznych obejmują informacje tekstowe [13], materiały wideo [14], zdjęcia [15] oraz komunikaty marketingowe [16]. Badania w tym obszarze mają charakter interdyscyplinarny z dużym udziałem obliczeniowej nauki o sieciach (ang. *computational network science*).

Wyzwania naukowe informatyki powiązane są m.in. z algorytmami wyboru wpływowych węzłów sieci [34][21], modelami przepływu informacji [17], rozwiązaniami analitycznymi, które umożliwiają identyfikację czynników wpływających na efektywność procesów [20], maksymalizacją wpływu [18][24] oraz analizą zjawisk związanych z propagacją zachowań [19]. Identyfikowane w tym obszarze zagadnienia generują wiele problemów analitycznych, obliczeniowych i charakteryzowane są jako problemy NP-trudne [25]. Zamierzona

inicjalizacja procesów propagacji informacji opiera się na wyborze zbioru początkowych węzłów z zakładanym dużym potencjałem wyzwiania kaskad informacyjnych (ang. *seed set selection problem, target set selection*), a następnie ich aktywacji poprzez przekazanie informacji do dalszego rozprzestrzeniania, co jest traktowane całościowo jako proces zasiewania (ang. *seeding*) [31]. Wyznaczenie zbioru początkowego wiąże się z wysokimi kosztami obliczeniowymi. Stosowane są tutaj rozwiązania przybliżone oparte na podejściu zachłannym [25] i jego rozszerzeniach [27], zastosowaniach technik eksploracji danych [26], czy modeli kombinatorycznych [22][23]. Często wykorzystywane są heurystyki na bazie miar centralności takich jak stopień wierzchołka, czy bliskość, dla których istnieje duży potencjał transmisyjny [28]. W badaniach własności procesów wykorzystuje się między innymi agentowe modele symulacyjne [29], które umożliwiają weryfikację nowych rozwiązań i wyznaczenie wpływu czynników takich jak charakterystyka sieci czy mechanizmów propagacji na dynamikę procesów rozprzestrzeniania informacji [30].

Większość wcześniejszych badań zakłada działania związane z wyborem węzłów początkowych w pojedynczym etapie i ich selekcja następuje tylko na początku procesu. Następnie informacja rozprzestrzenia się poprzez wykorzystanie naturalnych mechanizmów dyfuzyjnych bez dodatkowego wspomaganie. W modelowaniu propagacji informacji wykorzystuje się modele epidemiologiczne takie jak SIS, SIR i ich rozszerzenia [33], model niezależnych kaskad [25], modele gałązkowe [32], czy model liniowego progu aktywacji [25]. Poza rozwiązaniami jednoetapowymi powstały ostatnio prace bazujące na podejściu adaptacyjnym, w których wiedza zgromadzona podczas procesu wykorzystywana jest w celu poprawy jego parametrów [35]. Inne podejścia np. [37] uwzględniają równoległe procesy i interakcje między nimi, najczęściej w formie kaskad konkurujących [36].

Najnowsze badania w tym obszarze zorientowane są na optymalizację wykorzystania zasobów przeznaczonych na wybór węzłów początkowych oraz redukcję liczby selekcionowanych węzłów w tych samych segmentach sieci. Rozwiązania te oparte są przykładowo na wykorzystaniu mechanizmów głosowania przy zakładanym spadku wagi wraz ze zwiększaniem liczby aktywnych sąsiadów w otoczeniu [39]. Innym rozwiązaniem tego problemu jest algorytm *K-shell* oparty na identyfikacji centralnych węzłów, który zapewnia lepsze wyniki niż miary związane ze stopniami wierzchołków [40] oraz podejście zorientowane na ograniczenie selekcji w obrębie tych samych społeczności [38]. Rozwiązania te przyczyniają się do lepszego wykorzystania zbiorów początkowych i zwiększania zasięgu procesu przy takim samym koszcie, tj. liczbie węzłów aktywowanych w ramach zbioru początkowego.

Powyższe obszary stanowią jeden z kierunków badań w cyklu powiązanych tematycznie publikacji opartych na sekwencyjnym wyborze węzłów, który umożliwia lepsze wykorzystanie naturalnych procesów dyfuzyjnych [A1]. Wyniki badań związanych z selekcją sekwencyjną stanowiły podstawę realizowanych prac zorientowanych na opracowanie nowych algorytmów i metod inicjowania procesów dyfuzji informacji. Myślą przewodnią było dążenie do ograniczenia wykorzystania węzłów, które mogą być aktywowane w sposób naturalny oraz założenie, że może to się odbywać kosztem wydłużenia czasu trwania procesu w stosunku do rozwiązań jednoetapowych. W ramach badań weryfikowano różne podejścia i najlepsze efekty uzyskano w opracowanym algorytmie sekwencyjnym (ang. *sequential seeding*), którego specyfika odbiega od stosowanych dotychczas rozwiązań algorytmicznych. W opracowanych rozwiązaniach jest rozpatrywany kompromis między czasem trwania

procesu, a zasięgiem w ramach sieci złożonej, tj. liczbą aktywowanych węzłów. W badaniach zweryfikowano różne warianty algorytmu. Poza podstawową wersją, która zakłada sekwencjonowanie bezwarunkowe bez względu na stan procesu, zaproponowano również rozszerzenie w postaci algorytmu ze wzbudzaniem kaskad informacyjnych (ang. *revival mode*) oraz z buforowaniem (ang. *buffering mode*). W wersji rozszerzonej wykorzystanie dodatkowych aktywacji następuje dopiero w momencie wygaszenia procesu. Algorytmy zweryfikowano z udziałem symulacji agentowych przy zakładanym modelu rozprzestrzeniania informacji bazującym na niezależnych kaskadach (ang. *independent cascades model*) [25], na zbiorze rzeczywistych sieci złożonych przy różnych parametrach procesu i liczbie węzłów selekcjonowanych do zbioru początkowego. Porównano je z istniejącymi podejściami opartymi na miarach centralności, z algorytmem *VoteRank* [39] oraz z algorytmem zorientowanym na społeczności [38]. Wyniki potwierdziły poprawę rezultatów w stosunku do stosowanych wcześniej metod. Przykładowo w ponad 90% przebiegów symulacyjnych rozwiązania sekwencyjne zapewniają lepsze wyniki niż rozwiązania jednoetapowe oparte na miarach centralności takich jak *degree*, czy *closeness*. W badaniach przeprowadzono również analizy wpływu buforowania i wznawiania procesów na ich zasięg i czas realizacji. Algorytm z wzbudzaniem procesów zapewnia największy zasięg, ale jednocześnie powoduje wydłużenie czasu trwania procesu. Buforowanie umożliwia realizację procesu w czasie zbliżonym do algorytmu sekwencyjnego bez wznawiania i jest rozwiązaniem kompromisowym, które zapewnia zasięg niewiele mniejszy niż wersja z wznawianiem bez buforowania. Główny potencjał opracowanych rozwiązań bazuje na lepszym wykorzystaniu naturalnych procesów rozprzestrzeniania informacji i ograniczeniu zużycia zasobów na aktywację węzłów, które mogą być inicjowane w procesie naturalnym. Opóźniona selekcja węzłów w podejściu sekwencyjnym skutkuje poprawą zasięgu, ale jednocześnie przekłada się na wydłużenie czasu trwania procesu. Badania symulacyjne bazowały na różnej liczbie aktywacji w każdym etapie sekwencji, co umożliwiło wyznaczenie zależności pomiędzy zasięgiem, a wydłużeniem czasu trwania procesu. Największy zasięg uzyskano przy maksymalnej dekompozycji procesu i pojedynczej aktywacji w każdym kroku. Skutkowało to jednocześnie największym wydłużeniem procesu. Najlepsze wyniki w wymiarze zasięgu uzyskano dla algorytmu ze wzbudzaniem wygaszonych kaskad.

W podstawowym algorytmie sekwencyjnym w każdym kroku selekcja węzłów opiera się na rankingu wyznaczonym na podstawie miar sieciowych obliczonych na początku procesu. Może to prowadzić do sytuacji, gdy węzły o dużym potencjale dyfuzyjnym i przykładowo wysokim stopniu wierzchołka tracą potencjał w wyniku aktywacji w bezpośrednim otoczeniu, a z uwagi na wysoką pozycję w rankingu statycznym są selekcjonowane w kolejnych etapach sekwencji. W kolejnej pracy uwzględnionej w cyklu publikacji [A2] zaproponowano algorytm z wykorzystaniem miar efektywnych, przy wyznaczaniu których bierze się pod uwagę tylko nieaktywne w danym etapie procesu węzły sieci. Przykładowo przy wyznaczaniu dynamicznego stopnia wierzchołka DD (ang. *dynamic degree*) uwzględniane są tylko węzły z bezpośredniego sąsiedztwa, które nie zostały jeszcze aktywowane w procesie naturalnym lub podczas początkowego wyboru. W podobny sposób można wyznaczyć miary efektywne przy dynamicznym stopniu wierzchołka drugiego stopnia z uwzględnieniem liczby sąsiadów i sumarycznej liczby sąsiadujących węzłów wszystkich sąsiadów (DD2) (ang. *second level dynamic degree*) oraz analogicznie przy innych miarach sieciowych. Zastosowanie dynamicznych rankingów umożliwiło poprawę liczby aktywacji w obrębie sieci w ponad 80% wariantów konfiguracyjnych w takich samych warunkach. Przeliczanie miar sieciowych w poszczególnych etapach wyboru powoduje wzrost złożoności i potrzebę wykorzystania

dotychczasowych zasobów obliczeniowych. W badaniach analizowano czas przeliczania miar sieciowych i uzyskane dane pozwoliły na wyznaczenie zależności pomiędzy przyrostem zasięgu, a wzrostem kosztów obliczeniowych. W eksperymentach symulacyjnych badano zarówno różne parametry propagacji, jak i interwały pomiędzy przeliczaniem miar sieciowych oraz ich wpływ na zasięg procesu.

Nowym podejściem w realizowanych badaniach było rozpatrywanie dodatkowego oddziaływania na proces rozprzestrzeniania informacji po jego zainicjowaniu poprzez wspomagający wybór węzłów sieci i ich aktywacje (ang. *supporting seeding*) [A3]. W badaniach założono inicjowanie procesu z wykorzystaniem jednoetapowych metod opartych na miarach centralności i dodatkowe wsparcie poprzez aktywację wybranych węzłów sieci podczas trwania procesu. Opracowany algorytm na bieżąco wykorzystuje dane z procesu rozprzestrzeniania i dokonuje dodatkowej selekcji węzłów na podstawie ich miar centralności. W badaniach zweryfikowano działanie algorytmu w przestrzeni eksperymentalnej z uwzględnieniem różnych prawdopodobieństw propagacji. Założono poziomy wsparcia procesu poczynając od 10% do 100% zbioru początkowego. Badania pozwoliły na wyznaczenie efektywności przy różnych poziomach wsparcia reprezentowanego przez przyrost liczby dodatkowych aktywacji w sieci przypadających na oddziaływanie jednostkowe. Największy jednostkowy przyrost zasięgu zaobserwowano przy liczbie dodatkowo selekcjonowanych węzłów na poziomie 10% - 30% zbioru początkowego. Zwiększenie oddziaływania powyżej 30% zbioru początkowego przekładało się na gorsze efekty jednostkowe.

Dotychczasowe badania związane z inicjowaniem procesów dyfuzji informacji w sieciach złożonych bazują na homogenicznych zbiorach początkowych, które są selekcjonowane przy użyciu jednej z metod, na przykład rankingu na podstawie stopnia wierzchołka lub innej miary centralności. Podczas gdy każda z miar reprezentuje określone własności sieci, homogeniczny zbiór węzłów początkowych może prowadzić do nadreprezentacji węzłów o określonych charakterystykach. Jednocześnie wybór do zbioru początkowego elementów przykładowo z wysokim stopniem wierzchołka przyczynia się do selekcji węzłów które mogą być powiązane z innymi znajdującymi się w zbiorze początkowym [39]. Uwzględnianie ich w zbiorze początkowym nie jest konieczne, ponieważ istnieje duże prawdopodobieństwo, że mogą być one aktywowane w naturalnym procesie propagacyjnym. Przekłada się to na możliwość lepszego wykorzystania zasobów i skłania do poszukiwań innych rozwiązań. W badaniach [A4] podjęto próbę eliminacji tego problemu przy jednoczesnym zachowaniu efektywności procesu mierzonej liczbą aktywacji poprzez zastosowanie optymalizacyjnych modeli mieszanek. Przyjęto w nich wykorzystanie zbioru węzłów z różnym udziałem elementów o określonych własnościach np. 25% węzłów z rankingu generowanego na podstawie stopnia wierzchołka, 25% według miary pośrednictwa oraz 50% według miary bliskości. Zastosowanie mieszanek umożliwiło uzyskanie bardziej zróżnicowanego zbioru docelowego bez nadmiernej eksploatacji węzłów o określonych charakterystykach, a jednocześnie nie wpłynęło to na pogorszenie efektywności procesu. Rozwiązania mieszkankowe zweryfikowano z wykorzystaniem modelowania agentowego przy przestrzeni planów eksperymentalnych dla poszczególnych parametrów.

Poza badaniami sieci statycznych jednym z aktualnych kierunków badań są sieci dynamiczne, których reprezentacja z uwzględnieniem aspektów temporalnych zapewnia opis bardziej zbliżony do rzeczywistości niż ma to miejsce w strukturach statycznych [47]. Problem wyboru

zbioru początkowego jest utrudniony ze względu na zmienność właściwości sieci w czasie, które zaś wynikają ze zmiennej liczby węzłów i krawędzi. Utrudniona jest wtedy właściwa alokacja zasobów i inicjowanie procesów rozprzestrzeniania informacji. Większość dostępnych strategii selekcji węzłów początkowych bazuje na statycznych miarach centralności takich jak stopień wierzchołka, bliskość, czy współczynnik klasteryzacji, a w strategiach wyboru węzłów początkowych nie uwzględnia się parametrów dostępności węzłów. Problem ten występuje przykładowo w platformach społecznościowych, gdzie dla dynamiki procesu rozprzestrzeniania kluczowe są zalogowania użytkowników i aktywne korzystanie z systemu. Potencjał propagacyjny użytkownika z wysokim *degree*, ale z małą częstotliwością korzystania z systemu może być niewielki. W pracy [A5] zwrócono uwagę na możliwość wykorzystania w procesach dyfuzyjnych węzłów o dużej aktywności z gorszymi parametrami sieciowymi, których potencjał propagacyjny jest większy niż tych charakteryzujących się wysokimi miarami, ale zarazem mniejszą aktywnością. Według wiedzy Autorów jest to jedna z pierwszych prac, w której zaproponowano algorytm wyboru węzłów początkowych w sieciach dynamicznych. Przeprowadzone badania symulacyjne w środowisku agentowym dały możliwość określenia mechanizmów kompensacji miar centralności miarami aktywności. Zaproponowane rozwiązanie przyczynia się do równoważenia obciążenia w sieci poprzez mniejsze wykorzystanie kluczowych węzłów z wysokimi miarami centralności na rzecz węzłów mniej istotnych pod tym względem, lecz z większymi współczynnikami aktywności. Badania przeprowadzono z wykorzystaniem teoretycznych modeli sieciowych, takich jak model preferencyjnego dołączania Barabásiego-Alberta [44], model Erdősa-Rényi [46] oraz model Watts-Strogatza [45] bazujący na zjawisku małego świata (ang. *small-world phenomenon*). Wyniki pokazały, że w sieci Barabásiego-Alberta zbiór początkowy wyznaczony przy użyciu miar centralności może być skompensowany dwukrotnie większym zbiorem węzłów o mniejszych wartościach centralności, ale o większej aktywności w sieci. W artykule zaproponowano też wektorowy wybór węzłów początkowych, w którym przy selekcji początkowej jest wyznaczany wektor wartości maksymalnych dla każdej z uwzględnianych miar sieciowych, a następnie dla każdego węzła kandydującego określa się miarę euklidesową dystansu. Badania wykazały, że strategia wektorowa w sieciach dynamicznych przynosi efekty lepsze niż klasyczne metody oparte na pojedynczych miarach sieciowych.

Problematyka procesów rozprzestrzeniania treści w sieciach była rozpatrywana również w powiązaniu z procesami dyfuzji wiedzy [A6]. W ramach sieci społecznych następuje przepływ wiedzy eksperckiej, wiedzy o produktach, czy tej bezpośrednio powiązanej z działalnością organizacji. Korporacyjne platformy społecznościowe wykorzystywane są do jej propagacji i udoskonalania organizacji. Modelowanie zachodzących w nich zjawisk wymagało opracowania nowych rozwiązań. W przedstawionym modelu symulacyjnym uwzględniono charakterystyki poznawcze, kognitywistyczne i efekty emergentne powiązane z wspólnym nabywaniem wiedzy. Ujęto w nim rolę moderatora oraz przydziału ról eksperckich. Przedstawione algorytmy umożliwiają dodawanie nowych powiązań w sieci w celu wywarcia wpływu na kompetencje członków społeczności. W ramach badań opracowano model matematyczny i zweryfikowano go środowisku symulacyjnym z wykorzystaniem teoretycznych modeli sieci. Analizom podlegały procesy będące efektem zasilania poszczególnych węzłów lub segmentów sieci w wiedzę. Badania wykazały, że zbyt intensywne zasilanie powoduje niewykorzystanie potencjału alokowanych zasobów i przekłada się na obniżenie efektywności procesów propagacji w wyniku zapomnienia, czy niedopasowania wzorców kompetencji u nadawców i odbiorców.

Poza rozprzestrzenianiem informacji w sieciach jednowarstwowych aktualnym kierunkiem badawczym jest również dyfuzja informacji w sieciach wielowarstwowych [42]. W systemach rzeczywistych powiązania występują pomiędzy wieloma warstwami i mogą dotyczyć kontaktów prywatnych, zawodowych, czy komunikacji w platformach społecznościowych. Uwzględnienie tej specyfiki w badaniach daje możliwość lepszego odzwierciedlenia w modelach teoretycznych zjawisk zachodzących w rzeczywistości. Cykl powiązanych tematycznie publikacji obejmuje również badania z tego obszaru [A7]. Opracowany model uwzględnia specyfikę dyfuzji wewnątrzwarstwowej i międzywarstwowej oraz dekomponuje typowe modele wiedzy na podejście wielowarstwowe, podczas gdy dotychczasowe rozwiązania uwzględniają dyfuzję w obrębie jednej warstwy. Z organizacyjnego punktu widzenia modelowanie zmian kompetencji z podziałem na warstwy jest rozwiązaniem lepiej oddającym specyfikę rzeczywistych procesów przepływu wiedzy. Nowym elementem jest również rozpatrywanie procesów z udziałem dyfuzji poziomej w ramach jednej warstwy oraz pionowej między warstwami.

Realizowane badania sieci złożonych i platform społecznościowych wskazały na szereg problemów związanych z analizami i przetwarzaniem dużych zbiorów danych. Pojawiło się pytanie, czy dla potrzeb realizacji badań możliwe jest wykorzystanie próbek sieci zamiast kompletnych zbiorów danych. Dla potrzeby analizy charakterystyki procesów rozprzestrzeniania informacji ich własności powinny być zbliżone do charakterystyk sieci wyjściowej. Dostępne metody zapewniają generowanie reprezentatywnych próbek w postaci mniejszych sieci o zbliżonych własnościach do sieci pierwotnej, jednakże zazwyczaj z nadreprezentacją węzłów o najwyższych miarach centralności. W pracy [A8] zaproponowano podejście zrównoważone, w którym pozyskuje się próbki o zadanych rozkładach parametrów sieciowych. Zaproponowane algorytmy wykorzystują funkcję ewaluacyjną i ocenę jakości próbki w odniesieniu rozkładów miar w sieci pierwotnej. Uzyskane w ten sposób struktury na podstawie sieci wyjściowej mają zbliżone rozkłady parametrów wykorzystywanych przy selekcji węzłów początkowych i charakterystyki zbliżone do sieci podstawowej.

Kolejne prace w ramach cyklu powiązanych tematycznie publikacji nawiązują do klasyfikacji ACM w kategorii badań związanych zarówno z projektowaniem interakcji człowiek-komputer jak i badań użytkowników systemów informatycznych. W badaniach wykorzystano środowisko wirtualnego świata jako platformę do realizacji eksperymentów związanych z procesami dyfuzji informacji. Opracowane rozwiązania programistyczne umożliwiły integrację mechanizmów dyfuzji treści cyfrowych i monitoring procesów rozprzestrzeniania informacji. Badania rzeczywistych procesów dały podstawę do opracowania modelu, który odzwierciedla strukturę kaskad informacyjnych z uwzględnieniem poszczególnych poziomów aktywacji i ich charakterystyk w formie generacji. W artykule [A9] przedstawiono rozwiązanie oparte na podejściu generacyjnym, w którym rozpatrywano poziomy aktywacji z uwzględnieniem hierarchii powiązań, a nie tylko czynnika czasowego jak to ma miejsce w większości modeli analitycznych. Wymiar generacji dostarcza dodatkowych możliwości analitycznych o strukturze aktywacji węzłów, poziomach kaskad informacyjnych i ich parametrach ilościowych, co przekłada się na wiedzę o procesie i jego dynamice. W badaniach testujących wykorzystano dane z procesu na różnych etapach jego realizacji i analizie podlegała uzyskana dokładność modelu, która zwiększała się wraz z pozyskiwaniem danych. Nowym elementem jest mikroskopowe podejście analityczne oparte na hierarchicznej strukturze kaskad, a nie tylko na zagregowanych danych makroskopowych.

Celem dalszych badań była analiza wpływu parametrów treści cyfrowych oraz mechanizmów propagacji na dynamikę procesów rozprzestrzeniania informacji. W pracy [A10] rozpatrywano różne charakterystyki i mechanikę procesów rozprzestrzeniania informacji w platformie społecznościowej. Wyodrębniono niską barierę transmisji przekazu określaną jako *Low Resistance* (LR) oraz wysoką barierę transmisji określaną jako *High Resistance* (HR). Jako podstawę metodyczną analiz wykorzystano procesy gałązkowe oraz zaproponowane wcześniej podejście generacyjne dla modelowania kaskad informacyjnych. Na podstawie zarejestrowanych danych dla poszczególnych eksperymentów wyznaczono parametry: propagacji p , określający prawdopodobieństwo przekazywania treści, intensywności epidemii λ oraz progę epidemii (ETP) zdefiniowany jako $p * \lambda$. Zgodnie z [41], związek między ich wartościami określany jest jako podkrytyczny gdy $ETP < 1$, nadkrytyczny dla $ETP > 1$ oraz krytyczny dla $ETP = 1$. Eksperymenty dostarczyły wyników teoretycznych istotnych dla badania interakcji człowiek-komputer i powiązania technicznych mechanizmów propagacji treści z charakterystykami użytkowników. W analizach wprowadzono współczynniki konwersji i miary zaangażowania dla odzwierciedlenia charakterystyk ilościowych badanych zjawisk. Uzyskane wyniki wskazały na kluczowy udział czynników wizualnych przekazu, które miały większy wpływ na dynamikę transmisji niż parametry nadawców i odbiorców. W analizach uwzględniono wielowarstwowy charakter sieci społecznych z podziałem na warstwę komunikacyjną, transakcyjną oraz społeczną. Rezultaty wskazują na kluczową rolę fazy prezentacji produktu, która wzmacnia procesy dyfuzyjne oparte na komunikacji tekstowej. Odzwierciedla to procesy ze świata rzeczywistego, w których rozprzestrzenianie informacji może zarówno być powiązane z faktyczną prezentacją produktu jak i bazować tylko na opisie słownym. Analiza literatury wykazała brak badań tego typu realizowanych w środowisku cyfrowym. Wyniki wskazują na niską efektywność propagacji produktów rozprzestrzenianych w modelu *Low Resistance* z uwagi na brak relacji społecznych i oddziaływań na linii nadawca-odbiorca. Mimo potencjalnie dużej liczby transmisji konwersja i parametry zaangażowania były przy takiej transmisji relatywnie niskie. Uzyskane rozwiązania mogą stanowić podstawę do wyznaczania kompromisu pomiędzy transmisją przekazu z dużą dynamiką i mniejszą skutecznością na rzecz zwiększenia barier transmisyjnych i przełożenie tego na silniejsze relacje między odbiorcami a nadawcami. Podobną relację zaobserwowano dla działań związanych z premiowaniem użytkowników, co wpływało na wzrost liczby interakcji, ale nie prowadziło do zwiększenia efektów. Poza aspektem naukowym wyniki badań mogą zostać wykorzystane przez operatorów systemów internetowych przy projektowaniu mechanizmów powiązanych z propagacją produktów cyfrowych, których zwiększający się udział w rynku jest podkreślany w licznych opracowaniach [43].

Dla lepszego poznania zjawisk zachodzących w systemach społecznościowych, które mają wpływ dla procesów rozprzestrzeniania informacji zrealizowano badania interdyscyplinarne przedstawione w pracy [A11]. Ich celem było monitorowanie i analiza czynników wpływających na budowę więzi społecznych i oddziaływanie w platformach społecznościowych. Połączenie informatyki i socjologii umożliwiło zaplanowanie oraz implementację systemów pomiarowych w platformie online zorientowanych na monitorowanie i rejestrację zdarzeń związanych z nawiązywaniem relacji społecznych. Poza wynikami istotnymi z socjologicznego punktu widzenia w zakresie dołączania preferencyjnego (ang. *preferential attachment*) oraz generalizacji teorii związanych z formowaniem relacji projekt dostarczył interesujących wyników z perspektywy projektowania mechanizmów komunikacji i interakcji w platformach społecznościowych oraz

planowania scenariuszy, zdobywania poziomów zaawansowania oraz segmentacji użytkowników. Wiąże się to z grywalizacją (ang. *gamification*), doświadczeniami użytkownika (ang. *user experience*) oraz projektowaniem produktów wirtualnych, czy zasad subskrypcji. Realizacja projektu wpisuje się w nurt interdyscyplinarny, którego rola jest aktualnie podkreślana.

Kolejne badania w ramach modelowania i eksploatacji systemów społecznościowych były związane z sektorem produktów cyfrowych i ich rozprzestrzenianiem pomiędzy użytkownikami [A12]. Celem badań było modelowanie charakterystyk użytkowników realizujących transakcje związane z zakupami produktów cyfrowych w ramach platformy społecznościowej. Dla potrzeb budowy systemu wnioskowania wykorzystano rozmyty model ANFIS, który zintegrowano z bazami transakcyjnymi systemu. W badaniach przyjęto podejście makroskopowe z uwzględnieniem zagregowanych danych transakcyjnych bez analiz na poziomie struktur sieci społecznych. Wynikiem prac są dwa modele wnioskowania dotyczące zarówno charakterystyki użytkowników realizujących transakcje oraz także tych, którzy podejmują decyzje o transmisji obiektów wirtualnych i inicjują procesów ich propagacji w obrębie sieci.

W pracy [A13] zaproponowano modelowanie aktywności użytkowników w dwóch wymiarach czasowych. Dla celów wyznaczania parametrów aktywności przyjęto bitemporalny model reprezentacji czasu, w którym zdarzenia są zarejestrowane w wymiarze czasu rzeczywistego i wirtualnego. Ujęcie takie umożliwia reprezentację aktywności i zaangażowania użytkowników systemów społecznościowych.

Wkład w rozwój dyscypliny naukowej informatyka

Prowadzone prace badawcze dotyczą stosunkowo nowych obszarów, które dopiero od niedawna są przedmiotem intensywnych studiów naukowych. Rozwój platform społecznościowych i ewolucja systemów internetowych rodzi szereg wyzwań dla informatyki, związanych z ich modelowaniem, optymalizacją, eksploatacją, a zwłaszcza analizą pochodzących z nich danych (ang. *data science*). Liczne prace wskazują na potencjał badawczy systemów społecznościowych; notuje się znaczny udział liczby publikacji naukowych z obszaru informatyki w powiązanych tematach. Przykładowo dla hasła „*influence maximization*” w bazie SCOPUS zarejestrowanych jest 566 publikacji i 475 z nich jest klasyfikowanych w dyscyplinie informatyka, nawet dla hasła „*viral marketing*” na 791 publikacji 499 powiązanych jest z informatyką lub w szerszym kontekście dla „*social network analysis*” na 9.826 dokumentów 4.949 prac jest klasyfikowanych w informatyce, a tylko 2.847 w naukach społecznych.

Badania zrealizowane w ramach osiągnięcia naukowego mają istotny wpływ na rozwój informatyki; zgodnie z klasyfikacją ACM w zakresie *human-centered computing* ze szczególnym uwzględnieniem obszarów powiązanych z *social computing* takich jak *social networks*, *social media*, *social content sharing*, *collaborative content creation*, a także *human-computer design and evaluation methods* jak również *user studies*, *user models*, *web-based interaction* oraz *virtual reality*. Zaproponowane modele oraz wyniki badań eksperymentalnych przyczyniły się do lepszego poznania zjawisk zachodzących w systemach społecznościowych, a w efekcie do opracowania nowych metod zawiązanych z tą dziedziną. W szczególności, wartościowymi wynikami prac są metody modelowania i eksploatacji serwisów społecznościowych, które mogą być wykorzystane przy projektowaniu i analizie

systemów internetowych. Pozyskane wyniki stanowią również znaczący wkład do rozwijającej się aktualnie obliczeniowej nauki o sieciach (ang. *computational network science*). Jest to nowa dziedzina wspomagająca analizę złożonych systemów naturalnych i sztucznych oraz badanie relacji między różnymi komponentami sieci jak również między wieloma sieciami [1]. Jej podstawą jest podejście algorytmiczne, które przyczyniło się do ujednolicenia metod stosowanych w analizie sieci złożonych (ang. *complex networks*) w różnych dziedzinach.

Rozpatrywane w cyklu powiązanych tematycznie publikacji modelowanie przestrzeni wewnętrznej systemów społecznościowych bazuje na wykorzystaniu potencjału sieci społecznych i mechanizmów przepływu informacji. Realizowane w tym zakresie badania [A1] wprowadziły nowatorskie algorytmy zorientowane na sekwencyjną aktywację węzłów w sieci, co ma zasadnicze znaczenie przy inicjowaniu procesów dyfuzji informacji. Nowe metody były podstawą kompleksowych badań, które wykazały wzrost efektywności w porównaniu do algorytmów jednoetapowych stosowanych we wcześniejszych rozwiązaniach. Najwyższą efektywność i wykorzystanie potencjału naturalnych mechanizmów dyfuzji zapewniły algorytmy ze wzbudzaniem wygaszonych procesów w wersji podstawowej i z buforowaniem. Kontynuacją tych badań są rozwiązania oparte na dynamicznych rankingach [A2]. Badania nad nimi stanowią wkład w odkrywanie zależności pomiędzy złożonością obliczeniową, a efektywnością procesów dyfuzyjnych, w których kluczową rolę odgrywają struktury sieciowe. Prowadzone badania umożliwiły lepsze poznanie dynamiki zjawisk zachodzących w sieciach złożonych, a ich rezultaty mają potencjał do wykorzystania w modelowaniu systemów, dla działania których są istotne przepływy informacji. Dotyczy to zarówno badań w zakresie informatyki, systemów technicznych (np. rozprzestrzeniania się awarii) jak i modelowania zjawisk społecznych. Przedstawione rozwiązania związane z sekwencyjną selekcją węzłów inicjujących procesy rozprzestrzeniania informacji są podstawą metodyczną rozwiązań kompromisowych, w których można balansować pomiędzy końcowym zasięgiem, a czasem trwania procesu lub zasięgiem, a złożonością obliczeniową.

Uzyskane rezultaty mają duży potencjał rozwojowy i wskazują nowe kierunki badań sieci złożonych i procesów w nich zachodzących. Jednym z nich są prace związane z oddziaływaniem na procesy i zaproponowany algorytm zasiewu wspomagającego (ang. *supporting seeding*), w którym oddziałuje się na proces poprzez dodatkowe inicjowanie węzłów już w trakcie jego trwania [A3]. Jest to nowe podejście w tym obszarze badań, w którym dotychczasowe prace koncentrowały się na procesie głównym. W systemach rzeczywistych proces dyfuzji informacji nie jest zazwyczaj wyizolowany i pojawiają się w nim oddziaływania zewnętrzne związane między innymi z aktywacją dodatkowych węzłów. Szersze badania w tym obszarze są prowadzone aktualnie w ramach grantu Narodowego Centrum Nauki, którego jestem kierownikiem. Elementem nowości jest alternatywny dla homogenicznej selekcji węzłów w sieciach model mieszkankowy, który daje możliwość zróżnicowania zbioru początkowego i ustalenia różnych proporcji oraz charakterystyk węzłów inicjujących procesy dyfuzyjne [A4]. Realizowane badania stanowią wkład metodyczny w obszarze systemów i sieci dynamicznych ze strukturami zmiennymi w czasie. Obecnie jest to jeden z wiodących kierunków badań nad sieciami złożonymi. Zaproponowany w pracy [A5] kompensacyjny wybór węzłów początkowych jest jedną z pierwszych koncepcji, która rozpatruje selekcję węzłów inicjujących procesy w środowisku sieci dynamicznych.

Prowadzone badania wniosły nowe rozwiązania teoretyczne w zakresie kolaboracyjnego przetwarzania wiedzy, selekcji węzłów i ich zasilania dla zwiększenia efektywności procesów propagacji [A6][A7]. Zaproponowane modele przepływu i nabywania wiedzy w sieciach wprowadzają dynamiczne charakterystyki kompetencji i odzwierciadlają w systemie agentowym procesy rozprzestrzeniania. Oryginalność prowadzonych badań polegała na modelowaniu procesów dyfuzji wiedzy w ujęciu wielowarstwowym, w którym zaproponowano dyfuzję horyzontalną i wertykalną z mechanizmami samouczenia się i zapominania. Dotychczasowe modele bazowały na ujęciu jednowymiarowym z uwzględnieniem tylko jednym typem wiedzy. Opracowane rozwiązania agentowe umożliwiają analizę i śledzenie procesu rozprzestrzeniania i wzmacniania kompetencji w sieci, przydziału ról, a także usprawniają przepływy w grupach.

Oryginalnym rozwiązaniem jest podejście adaptacyjne do generowania próbek sieci, które ma na celu ich dopasowanie do zadanych rozkładów miar sieciowych wyznaczanych ze zbioru pełnego [A8]. Takie rozwiązanie umożliwia oddziaływanie na próbkę sieci poprzez dołączanie lub eliminację węzłów przy jednoczesnej ocenie i minimalizowaniu odległości charakterystyk próbki (rozkładów wybranych parametrów sieciowych) od charakterystyk zbioru pierwotnego z wykorzystaniem miary Kullbacka-Leiblera. Jedną z możliwych realizacji nowego podejścia jest zaproponowany algorytm *K-bins*, który na podstawie zadanych elementów składowych oceny próbki generuje docelowy zbiór danych przy minimalizacji odległości od charakterystyk zbioru pierwotnego. Ma to bardzo istotne znaczenie w dynamicznym próbkowaniu i predykcji procesów zachodzących w sieciach złożonych.

Badania związane z wykorzystaniem podejścia generacyjnego i procesów gałązkowych w analizie mechanizmów rozprzestrzeniania informacji skutkowały opracowaniem nowej metody analitycznej, która daje możliwość uwzględniania hierarchicznej struktury propagacji w sieci [A9]. Stosowane dotychczas metody zapewniają analizy w wymiarze czasu i zasięgu, a nie odzwierciedlają poziomów hierarchicznych związanych z kaskadami informacyjnymi. Analizy danych z poszczególnych poziomów dostarczają dodatkowej wiedzy o dynamice procesu i jego charakterystykach ilościowych. Zaproponowane podejście aproksymuje przebieg procesu do krzywej wzorcowej, której kształt i dokładność zależy od liczby kroków podlegających analizie.

Wyniki prezentowane w pracach [A10][A11][A12] dotyczące eksperymentów w ramach rzeczywistego systemu są istotne dla rozwijanych w ostatnim czasie badań związanych z tzw. grywalizacją (ang. *gamification*), które mają na celu uwzględnianie w projektach systemów mechanizmów pochodzących z gier, mających na celu zwiększenie zaangażowania i lojalności użytkowników. Dotyczyły one roli wskaźników statusu użytkownika w systemie, które rosną wraz z aktywnością i czasem korzystania z niego oraz zależności pomiędzy produktami udostępnianymi bezpłatnie (ang. *freemium model*), a produktami i usługami odpłatnymi. Badania odnoszą się też do specyfiki dynamicznie rozwijającego się sektora produktów wirtualnych, dla rozwoju których kluczowe są badania z obszarów informatyki związane z użytecznością, interakcją człowiek-komputer oraz projektowaniem systemów. Badania te były realizowane we współpracy z ośrodkami zagranicznymi i uznanymi ekspertami dziedzinowymi. W pracy [A10] za element nowości w obszarze *social content sharing* uznaje identyfikację wpływu mechaniki transmisji treści cyfrowych na procesy ich rozprzestrzeniania. Eksperymenty umożliwiły porównanie dynamiki procesów, w których odbiorca był poddawany oddziaływaniu wizualnemu transmitowanej treści z procesami

charakteryzującymi się oddziaływaniem jedynie werbalnym. Badania wpisują się w trend wykorzystania wirtualnych światów jako laboratoriów, w których można obserwować zjawiska trudne do monitorowania w świecie rzeczywistym oraz interakcję człowiek-komputer [2]. Praca [A11] wnosi rozwiązania w obszarach *user studies* oraz *virtual reality*. Wykazano w niej znaczący wpływ elementów projektowych systemu na dynamikę procesów komunikacyjnych, zwłaszcza w zakresie warstwy wizualnej systemu. Praca [A12] wskazuje nowe zastosowania modeli rozmytych w obszarze rozprzestrzeniania treści cyfrowych przez użytkowników (ang. *social content sharing*) i ich integracji z rzeczywistymi platformami społecznościowymi. Według wiedzy Autorów było to pierwsze rozwiązanie, które integruje modele wnioskowania rozmytego z danymi transakcyjnymi powiązаныmi z platformą dystrybucji produktów cyfrowych.

Analizy stanów wewnętrznych systemu związane są nie tylko ze strukturami sieci społecznych, ale też z wzorcami zachowań użytkowników. W badaniach [A13] zaproponowano modele zachowań z uwzględnieniem dwuwymiarowego modelu upływu czasu, w podziale na czas rzeczywisty i wirtualny. Zaproponowane podejście umożliwia monitorowanie zmian zachowań i predykcję stanów w oparciu o wprowadzoną funkcję dystansową, która wyznacza zależność między czasem rzeczywistym a wirtualnym. Spadek aktywności w obrębie systemu wpływa na wzrost dysproporcji, które są agregowane wraz z upływem czasu do pojedynczej miary dystansowej.

Reasumując, przedstawione w ramach cyklu powiązanych tematycznie publikacji pt. „Modelowanie przestrzeni wewnętrznej internetowych systemów społecznościowych” badania pozwoliły na lepsze poznanie złożonych systemów społecznościowych oraz wprowadziły do nauki rozwiązania algorytmiczne, metodyczne i analityczne na potrzeby modelowania zjawisk w nich zachodzących. Większość z nich ma szerokie możliwości zastosowań w modelowaniu i symulacji systemów, w których istotną rolę odgrywają sieci złożone i procesy rozprzestrzeniania informacji oraz przy rozwoju systemów wspomagających projektowanie oraz eksploatację systemów internetowych.

Prowadzone prace mają charakter międzynarodowy. Badania w ramach [A1-A5] prowadzono we współpracy z naukowcami z Politechniki Wrocławskiej oraz Rensselaer Polytechnic Institute, Troy, USA, prace [A8][A11] realizowano z naukowcami z Politechniki Wrocławskiej oraz Knowledge Media Research Center, Tübingen Niemcy, [A10] we współpracy z Game Research Lab, University of Tampere, Tampere, Finlandia oraz [A12] we współpracy z The University of Thessaly, Volos, Grecja.

Podsumowanie

Za najważniejsze osiągnięcia naukowe przedstawione w cyklu powiązanych tematycznie publikacji mające wkład w rozwój informatyki uznają:

- opracowanie algorytmów sekwencyjnego wyboru węzłów początkowych (i) w wariancie podstawowym, (ii) z wzbudzaniem procesów i (iii) z buforowaniem [A1], które wskazują nowy kierunek badań alternatywny do tradycyjnego podejścia jednoetapowego oraz zapewniają większy zasięg niż dotychczas stosowane algorytmy przy takiej samej liczbie aktywowanych węzłów początkowych,

- przeprowadzenie badań związanych z wpływem przyrostu zasięgu na wydłużenie czasu trwania procesów rozprzestrzeniania informacji [A1] oraz opracowanie algorytmów bazujących na rankingach dynamicznych i wyznaczanie zależności między zasięgiem procesu, a kosztami obliczeniowymi [A2],
- analizy związane z oddziaływaniem procesów wspomagających na proces pierwotny rozprzestrzeniania informacji, opracowanie algorytmów wspomagających [A3] oraz algorytmów wykorzystujących modele mieszkankowe do selekcji węzłów początkowych w dążeniu do uzyskania rozwiązań zrównoważonych [A4],
- opracowanie koncepcji kompensacyjnego wyboru węzłów początkowych w sieciach dynamicznych z uwzględnieniem parametrów aktywności oraz weryfikacja algorytmów w wersji podstawowej oraz wektorowej na teoretycznych modelach sieciowych [A5],
- opracowanie matematycznego modelu kolaboracyjnego rozprzestrzeniania wiedzy w sieci społecznej z uwzględnieniem detekcji społeczności oraz wektorów kompetencji [A6] i modelu wielowarstwowej dyfuzji w sieciach złożonych [A7],
- opracowanie metody próbkowania sieci oraz rozwiązań algorytmicznych opartych na minimalizacji odległości rozkładów miar sieciowych próbki od sieci pierwotnej z uwzględnieniem możliwości oddziaływania na trajektorię próbki w zakładanym przedziale wartości wyznaczanych przez zagregowane wartości dystansowe [A8],
- badania rzeczywistych procesów rozprzestrzeniania informacji oraz opracowanie i weryfikacja metody analizy generacyjnej procesów rozprzestrzeniania informacji w sieciach złożonych [A9], identyfikację wpływu mechaniki rozprzestrzeniania treści cyfrowych oraz roli czynników wizualnych na dynamikę i zasięg procesów rozprzestrzeniania informacji [A10],
- badania w zakresie wpływu mechanizmu dołączania preferencyjnego oraz statusu w systemach społecznościowych [A11], opracowanie modeli rozmytych dla ujęcia zachowań użytkowników powiązanych z produktami cyfrowymi [A12] oraz metod analitycznych opartych na czasie rzeczywistym i wirtualnym [A13].

Poza aspektami badawczymi i teoretycznymi uzyskane wyniki mają również charakter stosowany. Zasięwanie sekwencyjne wraz z rozszerzeniami stanowi nową meta-metodę, która może być wykorzystana w wielu obszarach praktycznych. W systemach technicznych sekwencyjne inicjowanie procesów propagacji informacji daje możliwość ograniczenia wykorzystania zasobów obliczeniowych oraz lepsze wykorzystanie przepustowości łączy. W przypadku komunikacji elektronicznej i działań marketingowych podejście sekwencyjne może przyczynić się do ograniczenia negatywnego oddziaływania na użytkowników w masowej komunikacji poprzez duże ilości przesyłanych wiadomości na rzecz lepszej alokacji zasobów i bazowania na naturalnych procesach dyfuzji informacji. Nie bez znaczenia jest też efektywność ekonomiczna podejścia sekwencyjnego i lepsze wykorzystanie dostępnych budżetów w działaniach marketingowych. Wyniki badań mogą znaleźć również zastosowanie w kontekście sytuacji kryzysowych oraz w propagowaniu informacji o zagrożeniach gdzie istotne jest uzyskanie jak największego zasięgu. Rezultaty stanowią wytyczne dla projektowania i eksploatacji systemów interaktywnych z uwzględnieniem specyfiki i oczekiwań ich użytkowników. Badania rzeczywistych procesów propagacji produktów wirtualnych, mechanizmów ich rozprzestrzeniania oraz charakterystyk użytkowników dostarczyły nowej wiedzy o funkcjonowaniu tych systemów. Eksperymenty związane z rolą awatarów i wizualnej reprezentacji użytkowników, statusów czy charakterystyk produktów wirtualnych, a także dynamiki relacji społecznych stanowią

podstawę do efektywnego projektowania i eksploatacji systemów [A10, A11, A12]. Potencjał gospodarczy oraz dynamiczny rozwój platform społecznościowych, sektora produktów wirtualnych czy gier społecznościowych realizowanych w modelu MMO (ang. *massively multiplayer online*) uzasadnia potrzebę prowadzenia dalszych prac badawczych w tym kierunku – kolejne etapy badań już są realizowane.

Pozostała działalność naukowo-badawcza

Tematyka pozostałych prac opublikowanych po uzyskaniu stopnia doktora nauk technicznych realizowanych w dyscyplinie informatyki w dużej mierze obejmuje badania systemów społecznościowych, ich modelowanie i zwiększanie ich efektywności. **Wykaz publikacji do których następują odwołania w tej części znajduje się w Załączniku 3.** W realizowanych badaniach związanych z przestrzenią wewnętrzną systemów społecznościowych analizowano algorytmy buforowania i wznawiania procesów propagacji informacji [13], badano wpływ parametrów odbiorców i nadawców przekazu na poziomie *diad* na efektywność procesu rozprzestrzeniania [22] z uwzględnieniem roli parametrów związanych z aktywnością w sieci. Kolejne badania w tym obszarze prowadzono w oparciu o sieci wielowarstwowe i zorientowano je na analizę procesów propagacji informacji w ujęciu wewnątrz i zewnątrz grupowym w sieciach rzeczywistych [23], a także wyznaczanie wartości progowych, których przekroczenie skutkuje gwałtownym wzrostem dynamiki procesów dyfuzyjnych w obrębie sieci złożonej [24]. Badania obejmowały również analizy struktur wielowarstwowych pod kątem występowania różnych społeczności w poszczególnych warstwach, co może prowadzić do trudności ze stosowaniem dotychczasowych algorytmów ich detekcji [21]. W realizowanych badaniach modelowano wieloetapowe mechanizmy przetwarzania danych w procesach dyfuzyjnych z podziałem na fazy zainteresowania przekazem, jego absorpcją i transmisją do innych węzłów. Opracowany model wieloetapowy umożliwił lepsze poznanie występujących zjawisk i rozszerza dostępne dotychczas rozwiązania o uwzględnienie charakterystyk przejść pomiędzy stanami węzłów [25]. W ramach badań rzeczywistych procesów wyznaczono zależności i funkcje przejść pomiędzy poszczególnymi etapami, które stanowią miarę efektywności mechanizmów wykorzystanych w procesie dystrybucji informacji w sieciach złożonych.

Kolejny obszar badawczy związany był z sektorem produktów wirtualnych i ich wykorzystaniem w internetowych systemach społecznościowych. Badania zrealizowane w tym obszarze obejmują system monitorowania transakcji i dynamicznej wyceny produktów cyfrowych dopasowanych do możliwości nabywczych użytkowników z udziałem modelowania rozmytego [19]. Przeprowadzone badania i eksperymenty umożliwiły wyznaczenie wpływu charakterystyk użytkowników na ich zachowania. Podsumowaniem badań w tym obszarze było opracowanie wpisu encyklopedycznego powiązanego z produktami wirtualnymi [78]. Badania związane z rozprzestrzenianiem wiedzy w sieciach społecznych były powiązane z opracowaniem założeń do systemu dynamicznego zarządzania kompetencjami [20]. Rozpatrywano także algorytmy rekomendujące oparte na automatach komórkowych w kontekście integracji danych z sieci społecznych i generowania rekomendacji na podstawie powiązań i relacji [27]. W ramach prac rozwinęto również metodę akwizycji wiedzy opartą na próbkowaniu [28].

Realizowane badania dotyczyły też użyteczności witryn internetowych i były prowadzone w kontekście zarządzania wiedzą dziedzinową w tym obszarze [39]. Wprowadzone

rozwiązania umożliwiają zbudowanie bazy wiedzy dziedzinowej, która ułatwia dobór metod oceny witryn internetowych. Elementem badań była selekcja kryteriów i doboru wag w procesie oceny projektów internetowych [34]. Na podstawie zaproponowanego rozwiązania zostały przeprowadzone badania użyteczności portali internetowych. Podstawą metodyczną proponowanych rozwiązań były wielokryterialne metody wspomagania decyzji, dla których zaproponowano rozwiązanie bazujące na opisie nieustrukturalizowanych problemów decyzyjnych. W pracy [40] przedstawiono metodę selekcji kryteriów i wyznaczania wag w procesie ewaluacji witryn internetowych. Elementem prac związanych z oceną witryn internetowych była również integracja w repozytorium ontologii dziedzinowych, które uwzględniają przestrzeń parametrów oceny systemów stosowanych w wiodących podejściach [36]. W ramach badań związanych z użytecznością witryn internetowych w pracy [37] przedstawiono rozwiązanie bazujące na wyznaczaniu kryteriów ewaluacyjnych o najwyższej istotności dla decydenta. Elementem badań była również budowa repozytorium wiedzy związanej z ewaluacją witryn internetowych [38].

Prace realizowane we wcześniejszym okresie nawiązywały do aktualnych trendów rozwojowych systemów internetowych i miały wpływ na kształtowanie moich późniejszych zainteresowań badawczych. Wraz z rozwojem mediów elektronicznych identyfikowano kierunki rozwoju systemów internetowych powiązane ze zwiększeniem udziału rozwiązań społecznościowych [72] oraz stopniową ewolucją systemów internetowych w tym kierunku [60]. Specyfika pierwszych platform społecznościowych wskazywała na możliwość wdrażania nowych rozwiązań analitycznych. Tematyka realizacji działań w sieciach społecznych była rozpatrywana również z uwzględnieniem czynników niepewności [75]. W ramach prac podejmowano tematy związane z modelowaniem przepływu informacji w sieciach społecznych i analizą dostępnych rozwiązań [60] a także rozwojem sektora mikrotransakcji oraz produktów wirtualnych w platformach społecznościowych i wirtualnych światach [61]. Jednym z tematów była identyfikacja czynników wpływających na powodzenie działań marketingowych w wirtualnych światach [62]. Zaproponowano integrację podejść afektywnych i pomiarów emocji w platformach społecznościowych [74], a także alternatywne podejście do klasyfikacji treści z udziałem selekcji kolaboracyjnej w aplikacjach realizowanych w nurcie Web 2.0 [81]. Przedstawiono również koncepcję przenoszenia paradygmatów Web 2.0 na grunt platform samorządowych [71] oraz rozpatrywano zagadnienia związane z projektowaniem architektury informacji w serwisach internetowych [76]. W ramach prac realizowano rozwiązania związane z monitorowaniem rozprzestrzeniania informacji marketingowych [58]. Jednym z efektów praktycznych był agentowy system monitorowania rynków elektronicznych [73]. Podjęto również temat klasyfikacji danych w platformach Web 2.0 [81] oraz zastosowania systemów wspomagania decyzji w handlu elektronicznym [77].

Poza przedstawioną działalnością naukową w dyscyplinie informatyka część mojego dorobku, szczególnie wcześniejszego, dotyczy zrównoważonego podejścia do marketingu elektronicznego, badań percepcyjnych, efektywności przekazu marketingowego, badań *eye trackingowych* oraz efektywności ekonomicznej. Publikacje z tych obszarów przedstawione w dalszej części autoreferatu wykazują na marginesie jako pozostały dorobek naukowy, nie związany bezpośrednio z osiągnięciem naukowym w dyscyplinie informatyka. Przedmiotem badań było między innymi zwiększanie konwersji i efektywności ekonomicznej w systemach internetowych [47] oraz wykorzystanie *eye trackingu* do wyznaczania charakterystyk związanych z obszarami zainteresowań oraz podziałem na elementy marketingowe

i redakcyjne w obrębie witryn internetowych [41]. W pracy [54] przeprowadzono z udziałem *eye trackingu* badania wpływu charakterystyk przekazu marketingowego na zachowania użytkowników. *Eye tracking* wykorzystano również przy analizach zachowań użytkowników systemu mobilnego w różnych grupach wiekowych [10]. W badaniach [45] zaproponowano dynamiczną zmianę poziomów oddziaływania na odbiorcę podczas powtórnych ekspozycji przekazu marketingowego, a w pracy [46] ograniczenie negatywnego wpływu przekazu o dużej intensywności. W tym nurcie opracowano model przekazu zorientowany na zmniejszanie inwazyjności komponentów reklamowych [48], przedstawiono założenia doboru poziomów oddziaływania perswazyjnego na użytkownika [43] oraz model optymalizacji działań reklamowych [9]. W pracy [56] zaproponowano metodę wieloetapowego zwiększania efektywności działań marketingowych, a następnie skoncentrowano się na interakcji pomiędzy komponentami graficznymi oraz elementami tekstowymi [44]. W badaniach zwrócono również uwagę na występowanie negatywnych efektów związanych z premiowaniem użytkowników za zaangażowanie w działania marketingowe [26].

W artykule [42] przedstawiono wyniki badań wzorców aktywności mózgu podczas przetwarzania treści marketingowych, opracowano wieloetapowy model zwiększania efektywności w realizacji działań marketingowych [51] oraz przedstawiono założenia optymalizacji kampanii reklamowych [50] i zastosowania metody COMET [49]. W pracy [64] zaproponowano ograniczenie nadmiernego doprecyzowania parametrów przekazu w targetowaniu behawioralnym. W ramach badań zaproponowano wykorzystanie wiedzy kolektywnej w procesie optymalizacji przekazu marketingowego w obrębie witryn internetowych [33]. Pozostałe publikacje obejmują planowanie działań marketingowych z uwzględnieniem czynników niepewności [82] oraz komunikację marketingową w aplikacjach Web 2.0 [57] i zachowania powiązane z przekazywaniem prezentów w systemie elektronicznym [12] (ang. *gifting behavior, gift economy*). Tematyka prac obejmowała również planowanie kampanii reklamowych w środowisku interaktywnym [63]. Media elektroniczne rozpatrywano jako środowisko z dużą dynamiką zmian [79] oraz opracowano założenia systemu zwiększania efektywności działań marketingowych w systemach wyszukiwawczych [67]. Zaproponowano rozwiązania adaptacyjne dla modelowania struktury przekazu reklamowego [80] oraz możliwości zastosowania analizy czynnikowej w personalizacji przekazu interaktywnego [66]. Ukazano również koncepcję zastosowania metod rekomendujących w systemie personalizacji produkcji [68], procedurę optymalizacji struktury treści marketingowych z udziałem analizy wielowymiarowej [69] oraz identyfikację elementów wpływających na efektywność działań marketingowych [70].

W ramach pozostałych prac przedmiotem badań były charakterystyki metod wspomagania decyzji stosowanych w zarządzaniu. W pracy [53] opracowano zbiór charakterystyk dla metod wielokryterialnego wspomagania decyzji, w [52] zaproponowano transformację zadań decyzyjnych. Zarządzanie wiedzą w obszarze metod wspomagania decyzji z uwzględnieniem charakterystyki problemów decyzyjnych przedstawiono w pracy [29]. W kolejnych badaniach zastosowano ontologie do reprezentacji wiedzy związanej ze specyfiką wielokryterialnych metod wspomagania decyzji [30] oraz przeprowadzono badania przy niedokładnych ocenach ekspertów [55]. Metody wielokryterialnego wspomagania decyzji zostały też wykorzystane do oceny efektywności witryn internetowych [31], systemów aukcyjnych [35], przy ocenie potencjału ekonomicznego lokalizacji źródeł energii [11], a także w inteligentnych łańcuchach dostaw [32].

Podsumowanie dorobku naukowego

Jestem autorem oraz współautorem 82 prac opublikowanych po uzyskaniu stopnia doktora. Dwadzieścia siedem z nich to prace samodzielne, dwanaście prac zostało opublikowanych w czasopismach z listy Journal Citation Report (JCR) posiadających współczynnik Impact Factor. W Tabeli 2 przedstawiono wskaźniki bibliometryczne całości dorobku po uzyskaniu stopnia doktora (na podstawie wykazu publikacji przedstawionych w **Załączniku 3**).

Tabela 2. Wskaźniki całości dorobku naukowego po uzyskaniu stopnia doktora

Nazwa wskaźnika	Wartość
Liczba publikacji w czasopismach z Impact Factorem	12
Sumaryczny wskaźnik Impact Factor	24.34
Index H według Web of Science	6
Index H według Google Scholar	10
Liczba cytowań Web of Science bez autocytowań	51
Liczba cytowań Google Scholar bez autocytowań	97
Liczba publikacji indeksowanych w bazie Web of Science	46
Liczba publikacji na konferencjach międzynarodowych	37
Liczba publikacji z listy B MNiSW	16
Liczba rozdziałów w monografiach	7
Sumaryczna liczba punktów MNiSW	1091
Liczba punktów MNiSW według udziału	524.8

Dorobek dydaktyczny, popularyzatorski i informacje o współpracy międzynarodowej przedstawiono w **Załączniku 4**.

Bibliografia

1. Hexmoor, H. (2014). *Computational Network Science: An Algorithmic Approach*. Morgan Kaufmann.
2. Bainbridge, W. S. (2007). The scientific research potential of virtual worlds. *Science*, 317(5837), 472-476.
3. Chu, S. C. (2011). Viral advertising in social media: Participation in Facebook groups and responses among college-aged users. *Journal of Interactive Advertising*, 12(1), 30-43.
4. Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.
5. Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295-298.
6. Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *Information Fusion*, 28, 45-59.
7. Giri, K. J., & Lone, T. A. (2014). Big Data-Overview and Challenges. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(6).
8. Russell, M. A. (2013). *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More*. O'Reilly Media, Inc.
9. Brown, J. R., van Dam, A., Earnshaw, R., Encarnação, J., Guedj, R., Preece, J., ... & Vince, J. (1999). Human-centered computing, online communities, and virtual environments. *IEEE Computer Graphics and Applications*, 19(6), 70-74.
10. Raice, S. (2012). Facebook sets historic IPO. *The Wall Street Journal*, (p. A11).
11. Chang, H. C. (2010). A new perspective on Twitter hashtag use: Diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1-4.
12. Hoogendoorn, M., Treur, J., van der Wal, C. N., & van Wissen, A. (2010, August). An agent-based model for the interplay of information and emotion in social diffusion. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology- Volume 02* (pp. 439-444). IEEE Computer Society.
13. Bastos, M. T., Raimundo, R. L. G., & Travitzki, R. (2013). Gatekeeping Twitter: message diffusion in political hashtags. *Media, Culture & Society*, 35(2), 260-270.
14. Susarla, A., Oh, J. H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, 23(1), 23-41.
15. Wei, X., Valler, N. C., Prakash, B. A., Neamtiu, I., Faloutsos, M., & Faloutsos, C. (2013). Competing memes propagation on networks: A network science perspective. *IEEE Journal on Selected Areas in Communications*, 31(6), 1049-1060.
16. Domingos, P. (2005). Mining social networks for viral marketing. *IEEE Intelligent Systems*, 20(1), 80-82.
17. Kossinets, G., Kleinberg, J., & Watts, D. (2008). The structure of information pathways in a social communication network. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 435-443). ACM.
18. Mossel, E., & Roch, S. (2007). On the submodularity of influence in social networks. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing* (pp. 128-134). ACM.
19. Centola, D. (2010). The spread of behavior in an online social network experiment. *science*, 329(5996), 1194-1197.
20. Ho, J. Y., & Dempsey, M. (2010). Viral marketing: Motivations to forward online content. *Journal of Business research*, 63(9), 1000-1006.
21. Kiss, C., & Bichler, M. (2008). Identification of influencers—measuring influence in customer networks. *Decision Support Systems*, 46(1), 233-253.

22. Ackerman, E., Ben-Zwi, O., & Wolfowitz, G. (2010). Combinatorial model and bounds for target set selection. *Theoretical Computer Science*, 411(44-46), 4017-4022.
23. Chiang, C. Y., Huang, L. H., Li, B. J., Wu, J., & Yeh, H. G. (2013). Some results on the target set selection problem. *Journal of Combinatorial Optimization*, 25(4), 702-715.
24. Morone, F., & Makse, H. A. (2015). Influence maximization in complex networks through optimal percolation. *Nature*, 524(7563), 65-68.
25. Kempe, D., Kleinberg, J., & Tardos, É. (2003, August). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 137-146). ACM.
26. Chen, W., Yuan, Y., & Zhang, L. (2010, December). Scalable influence maximization in social networks under the linear threshold model. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on* (pp. 88-97). IEEE.
27. Goyal, A., Lu, W., & Lakshmanan, L. V. (2011, March). Celf++: optimizing the greedy algorithm for influence maximization in social networks. In *Proceedings of the 20th international conference companion on World wide web* (pp. 47-48). ACM.
28. Chen, W., Wang, Y., & Yang, S. (2009, June). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 199-208). ACM.
29. Hui, C., Goldberg, M., Magdon-Ismael, M., & Wallace, W. A. (2010). Simulating the diffusion of information: An agent-based modeling approach. *International Journal of Agent Technologies and Systems (IJATS)*, 2(3), 31-46.
30. Bampo, M., Ewing, M. T., Mather, D. R., Stewart, D., & Wallace, M. (2008). The effects of the social structure of digital networks on viral marketing performance. *Information systems research*, 19(3), 273-290.
31. Hinz, O., Skiera, B., Barrot, C., & Becker, J. U. (2011). Seeding strategies for viral marketing: An empirical comparison. *Journal of Marketing*, 75(6), 55-71.
32. Van der Lans, R., Van Bruggen, G., Eliashberg, J., & Wierenga, B. (2010). A viral branching model for predicting the spread of electronic word of mouth. *Marketing Science*, 29(2), 348-365.
33. Wang, Y., Cao, J., Jin, Z., Zhang, H., & Sun, G. Q. (2013). Impact of media coverage on epidemic spreading in complex networks. *Physica A: Statistical Mechanics and its Applications*, 392(23), 5824-5835.
34. Kimura, M., Saito, K., Nakano, R., & Motoda, H. (2010). Extracting influential nodes on a social network for information diffusion. *Data Mining and Knowledge Discovery*, 20(1), 70-97.
35. Seeman, L., & Singer, Y. (2013, October). Adaptive seeding in social networks. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on* (pp. 459-468). IEEE.
36. Wei, X., Valler, N. C., Prakash, B. A., Neamtiu, I., Faloutsos, M., & Faloutsos, C. (2013). Competing memes propagation on networks: A network science perspective. *IEEE Journal on Selected Areas in Communications*, 31(6), 1049-1060.
37. Stanoev, A., Trpevski, D., & Kocarev, L. (2014). Modeling the spread of multiple concurrent contagions on networks. *PloS one*, 9(6), e95669.
38. Zhang, X., Zhu, J., Wang, Q., & Zhao, H. (2013). Identifying influential nodes in complex networks with community structure. *Knowledge-Based Systems*, 42, 74-84.
39. Zhang, J. X., Duan-Bing Chen, Q. D., & Zhao, Z. D. (2016). Identifying a set of influential spreaders in complex networks. *Scientific reports*, 6.
40. Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature physics*, 6(11), 888-893.
41. Becker, N. G. (1989). *Analysis of infectious disease data* (Vol. 33). CRC Press.
42. Salehi, M., Sharma, R., Marzolla, M., Magnani, M., Siyari, P., & Montesi, D. (2015). Spreading processes in multilayer networks. *IEEE Transactions on Network Science and Engineering*, 2(2), 65-83.
43. Huang, E. (2012). Online experiences and virtual goods purchase intention. *Internet Research*, 22(3), 252-274.

44. Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512.
45. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.
46. Erdős, P., Rényi, A., (1960). On the evolution of random graphs, *Evolution*, vol. 5, no. 1, pp. 17-61, 1960
47. Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics reports*, 519(3), 97-125.