

Prof. dr hab. inż. Andrzej Brykański
Katedra Zastosowań Informatyki
Wydział Elektryczny
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
Al. Piastów 17, 70-310 Szczecin
Andrzej.Brykański@zut.edu.pl

RECENZJA ROZPRAWY DOKTORSKIEJ

"Metody obliczeń rozproszonych w stochastycznym modelowaniu tkanek",
autorstwa mgr inż. Artura Sławomira Krupy

Promotor w przewodzie doktorskim: dr hab. inż. Bartosz Sawicki

Promotor pomocniczy: dr inż. Tomasz Leś

Recenzję wykonano na zlecenie Dziekana Wydziału Elektrycznego Politechniki Warszawskiej
(pismo przewodnie z dnia 29.03.2019 r.)

1. Ocena tematu i tezy pracy

Od kilkudziesięciu lat obserwuje się znaczący wzrost zainteresowania środowisk naukowych, medycznych oraz opinii społecznej zagadnieniami wpływu pola elektromagnetycznego na organizmy żywe, a w szczególności, od ponad ćwierćwiecza zainteresowania komputerowym modelowaniem zjawisk zachodzących w tkankach, czego wymiernym dowodem są liczne publikacje naukowe, które regularnie pojawiają się na światowym rynku wydawniczym. Jest to całkowicie zrozumiałe w sytuacji, w której w przestrzeni publicznej coraz częściej pojawia się pojęcie "elektrosmogu" oraz, gdy niezbędnym elementem projektowania nowych urządzeń emitujących pole elektromagnetyczne są gruntowne badania symulacyjne i eksperymentalne. Wpływ pola elektromagnetycznego na organizmy leży w obszarze zainteresowania, nie tylko różnych działów medycyny, szczególnie medycyny pracy, lecz przede wszystkim producentów sprzętu elektrycznego do zastosowań medycznych, przemysłowych lub codziennych zastosowań w warunkach domowych.

Badania symulacyjne, dotyczące procesów bioelektromagnetycznych należy uznać za wyjątkowo trudne, głównie za sprawą skomplikowanych struktur analizowanych tkanek, ich zmiennością oraz trudnymi do przewidzenia wartościami parametrów "materiałowych" tkanek (np. wartościami ich konduktywności). Wszystko to sprawia, iż symulacje komputerowe oparte o analizy deterministyczne, typowe dla różnych urządzeń technicznych, gdzie znane lub przynajmniej przewidywalne są parametry materiałowe, nie sprawdzają się w przypadku symulacji bioelektromagnetycznych. W takiej sytuacji, modelowanie tkanek winno mieć charakter modelowania stochastycznego, opartego o dane statystyczne. Niestety, takie podejście wymaga posiadania parametrów materiałowych, zawierających dane w ujęciu statystycznym (np. histogramy). Oznacza to konieczność akwizycji licznych danych, tak by miały one rzeczywistą wartość statystyczną. Jest to niezwykle trudne w przypadku, gdy materiał badawczy dotyczy tkanek ludzkich. Wynika to z ograniczeń liczebności materiału, który byłby cenny i z klinicznego i ze statystycznego punktu widzenia, przy jednoczesnych, wymagających warunkowaniach narzucanych przez towarzystwa bioetyczne i właściwe, etyczne komisje szpitalne.

Intensywnie rozwijany aparat matematyczny, służący praktycznym obliczeniom numerycznym, znajduje wsparcie w niezwykle dynamicznym rozwoju elektronicznej techniki obliczeniowej, zarówno na poziomie pojedynczych jednostek obliczeniowych, jak i "farm" komputerów, składających się na potężne narzędzia symulacyjne, działające w ramach tzw. maszyn wirtualnych w chmurach obliczeniowych (cloud computing). Zaletą obliczeń w chmurze tzw.

publicznej (*public cloud*) jest możliwość dysponowania ogromną mocą obliczeniową, bez konieczności posiadania własnej, złożonej infrastruktury komputerowo-sieciowej, często bez konieczności posiadania własnego, kosztownego, specjalistycznego oprogramowania, przy jednoczesnie bardzo atrakcyjnych kosztach korzystania z przeróżnych usług oferowanych przez dostawców chmury, zapewnieniu odpowiedniej wydajności, niezbędnych zabezpieczeń, chroniących składowane dane przed ich utratą w razie awarii i przed nieautoryzowanym dostępem, skalowalności etc.

Dzięki ofercie dostawców chmur obliczeniowych upowszechniło się korzystanie z dysków wirtualnych (eDyski), współdzielona praca z dokumentami typu arkusze kalkulacyjne, czy współtworzenie baz danych w chmurze; całkowicie realne się stało korzystanie z mediów cyfrowych, wykorzystywanie możliwości Internetu rzeczy (IoT), a także wykonywanie złożonych symulacji numerycznych, wymagających dużej mocy obliczeniowych i wiele innych.

W recenzowanej rozprawie doktorskiej, której tematyka mieści się w szerszym i bardzo aktualnym nurcie prac nad efektywnym, komputerowym modelowaniem zjawisk zachodzących w tkankach, podjęto próbę zastosowania obliczeń rozproszonych, wykorzystując do tego celu, jako narzędzie, platformę *Microsoft Azure*, oferującą zestaw różnych usług w chmurze. Należy przy tym zauważyć, że od strony koncepcyjnej, praca ma charakter szerszy i jej wyniki są możliwe do zaimplementowania również przy innych rozwiązaniach chmur obliczeniowych (np. dostawcy: *Amazon Web Services*, *Google Cloud Platform*, czy prawdopodobnie również rodzimy *Atman Cloud*).

Obliczenia numeryczne opierające się na materiale statystycznym (w tym przypadku stochastyczne modelowanie tkanek) posiadają znacznie większe wymagania, dotyczące wielkości mocy obliczeniowej, niż analogiczne analizy deterministyczne. Wykorzystanie skalowalnego systemu *cloud computing* może bez wątpienia przynieść tu spodziewane efekty w postaci bardziej wiarygodnych wyników symulacji komputerowych, przy jednoczesnym ograniczeniu czasu obliczeń numerycznych oraz stosunkowo niewysokich kosztach za komercyjne korzystanie z chmury.

Recenzowana praca jest wielowątkowa i interdyscyplinarna. Tematyka rozprawy leży na pograniczu trzech obszarów, zlokalizowanych w dziedzinie nauk technicznych: subdyscypliny informatyka (w ramach aktualnej dyscypliny naukowej „informatyka techniczna i telekomunikacja”), subdyscypliny elektrotechnika (w ramach aktualnej dyscypliny naukowej „automatyka, elektronika i elektrotechnika”) oraz bioinżynierii.

Główny obszar zainteresowań doktoranta dotyczy obecnej w informatyce technicznej, technologii *cloud computing*, a oś recenzowanej pracy stanowią nowoczesne i ciągle rozwijane narzędzia informatyczne.

Sformułowanie zasadniczego celu pracy:

Opracowanie metod i narzędzi do analizy numerycznej wykorzystujących chmurę obliczeniową do modelowania problemów bioelektromagnetycznych uwzględniając stochastyczny charakter parametrów materiałowych tkanek.

zostało przez Autora rozprawy poprzedzone gruntownymi studiami literaturowymi. Konstruktwna ocena zawartości przytoczonej, adekwatnej bibliografii, w większości pochodzącej z ostatniej dekady, umożliwiła Doktorantowi doskonałe rozeznanie w swoim obszarze badawczym oraz wyszukanie problemów otwartych, istotnych z punktu widzenia nauki, z którymi w przewodzie doktorskim postanowił się zmierzyć.

Stwierdzam, że zarówno sformułowany przez Doktoranta główny cel pracy, jak i przyjęte założenia oraz szczegółowe cele badawcze wraz z metodami ich realizacji są właściwe dla rozpraw doktorskich, przygotowywanych w naukach technicznych i w konfrontacji z aktualnym stanem wiedzy w dyscyplinie „Informatyka” wypadają pozytywnie.

2. Ocena merytoryczna pracy

Praca jest napisana na 129 stronach, z czego jej część zasadnicza z załącznikami obejmuje 113 stron. Zawiera ona obok rozdziału wstępnego, wprowadzającego do tematyki badań, omawiającego wybrane, adekwatne pozycje literaturowe oraz prezentującego cel badawczy, trzy rozdziały merytoryczne zawierające osiągnięcia badawcze Doktoranta, rozdział podsumowujący, a także bibliografię, liczącą 87 pozycji - głównie oryginalnych.

Praca zawiera także dwa załączniki, z których drugi, w mojej opinii powinien się znaleźć z zasadniczej części pracy, jako jeden z rozdziałów merytorycznych. Zawiera on istotne treści, m.in. dotyczące wdrożenia przygotowanego przez Doktoranta systemu obliczeniowego na bazie rozwiązania chmurowego *Microsoft Azure*. Załącznik ten jest autorskim opracowaniem mgr inż. Artura Sławomira Krupy, świadczącym o bardzo dobrym opanowaniu przez niego warsztatu naukowego i wdrożeniowego, głównie w zakresie zastosowań narzędzi *cloud computing* (m.in. wirtualizacji), wykorzystywanych w praktyce do zrównoleglonych symulacji komputerowych z danymi o charakterze stochastycznym.

Należy zauważyć, że około połowę pozycji bibliograficznych stanowią publikacje, które ukazywały się w ciągu ostatnich dziesięciu lat, co świadczy o aktualności podjętego przez Doktoranta przedsięwzięcia badawczego. Uwagę zwraca fakt, iż 8 pozycji literaturowych stanowią artykuły w czasopismach naukowych i materiałach konferencji międzynarodowych, których Doktorant jest współautorem. Wśród czasopism są reprezentowane: Rozprawy Elektrotechniczne (x2) oraz Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Środowiska (x1). Wśród konferencji występują głównie te, które dotyczą aspektów obliczeniowych z wykorzystaniem najnowszych narzędzi informatycznych. Zamieszczone w Bibliografii publikacje Doktoranta są ściśle powiązane z tematyką omawianą w poszczególnych, merytorycznych rozdziałach dysertacji.

Zawartość dysertacji uzupełniają takie elementy jak: pomocny wykaz ważniejszych oznaczeń i terminów oraz streszczenie w języku polskim, a także wymagane przepisami streszczenie angielskie.

Rozprawa jest bogato ilustrowana. Zawiera 62 rysunki, 9 tabel i cechuje ją na ogół dobra szata graficzna.

Rozdział pierwszy zawiera bogaty przegląd literatury przedmiotu, przy czym przytoczone publikacje, są ułożone w różnych dyscyplinach naukowych. Jest sprawą zrozumiałą, że dyscypliną wiodącą wśród tych publikacji jest „Informatyka” (algorytmizacja i obliczenia komputerowe, obliczenia typu *cloud computing*), jako służebna wobec innych dyscyplin technicznych i nietechnicznych, w szczególności ułożonych w obszarze elektromagnetyzmu.

Analiza literatury, dotyczącej wykorzystania narzędzi informatycznych w zagadnieniach bioelektromagnetyzmu, utwierdziła Doktoranta w przekonaniu, że brak jest reprezentacji publikacyjnej, która w miarę kompleksowo omawia zagadnienia obliczeń rozproszonych w stochastycznym modelowaniu tkanek i efektywnie wykorzystuje w tym celu, informatyczne rozwiązania typu *cloud computing*.

Rozdział drugi należy, obok dwu kolejnych, do rozdziałów kluczowych. Zawiera on istotne informacje dotyczące danych wejściowych, niezbędnych do stochastycznych analiz komputerowych tkanek. Doktorant zwraca uwagę na istotne fakty, które należy wziąć pod uwagę, by

uzyskać w miarę adekwatny model obliczeniowy. Podkreśla konieczność podejścia stochastycznego, które jest podyktowane złożonością badanych tkanek, wynikającą z ich skomplikowanego, niejednorodnego i nie zawsze powtarzalnego układu w sensie geometrycznym oraz zmienności parametrów „materiałowych” w różnych, zwykle niepowtarzalnych warunkach akwizycji danych. Omawia zasady tworzenia modelu statystycznego tkanek oraz sposoby praktycznej akwizycji danych na przykładzie materiału niejednorodnego (tkanka mięsno-tłuszczowa) i jednorodnego (ser żółty). Ten aspekt należy uznać za bardzo istotny w sytuacji, gdy nie można liczyć na odpowiednie dane uzyskane w warunkach klinicznych, gdzie są one zwykle reprezentowane nielicznymi przypadkami, zabranymi od różnych pacjentów, często na przestrzeni wielu lat i w różnych, często odmiennych warunkach. Doświadczenie pokazuje, że same opisy danych klinicznych są różne, nawet jeśli pochodzą one z tych samych laboratoriów. Jeśli nawet pewne elementy pomiarów w warunkach klinicznych są standaryzowane i odpowiadają wymogom w procesie leczenia, to zwykle w wyniku tych pomiarów nie uzyskuje się już innych, dodatkowych danych, które są niezbędne do polowych, stochastycznych analiz komputerowych. Wspomniane problemy stały się, jak informuje Doktorant: „inspiracją do własnych badań w zakresie pomiarowym oraz opracowania technologii pozwalającej na rozwiązanie kwestii liczby sparametryzowanych próbek tkanek biologicznych”. W dalszej części pracy (rozdział czwarty) zebrane doświadczenia opisane w rozdziale drugim, pozwoliły na wykorzystanie niektórych danych uzyskanych przez Doktoranta bezpośrednio na własnym stanowisku pomiarowym lub właściwego wykorzystania nielicznych (niestety) danych pochodzenia klinicznego z baz ogólnodostępnych (np. WHO), tak by można je było uznać za zadowalające z punktu widzenia statystycznego.

Oryginalne lub oryginalne częściowo treści są zawarte w rozdziale trzecim, w którym Autor dysertacji scharakteryzował m.in. wybrane, reprezentatywne, opisane w literaturze, te istniejące i współcześnie użytkowane oraz te, które mają już znaczenie historyczne, systemy i technologie obliczeń rozproszonych. Przegląd literaturowy, zawarty tej części pracy wraz z komentarzem Doktoranta ułatwiają lekturę dalszej części dysertacji. Miał on na celu zbadanie obecnie dostępnych rozwiązań obliczeń rozproszonych.

W tym samym rozdziale zawarte są dokonane przez Autora dysertacji bardzo ważne, wstępne testy systemów dystrybucji zadań, które pokazują m.in. wpływ liczby rdzeni (od 1 do 21) obecnych w klastrze na czas wielokrotnie powtarzanych obliczeń optymalizacyjnych o charakterze deterministycznym. Przykładowo, zastosowanie klastra (otwarta platforma programistyczna *Apache Hadoop*) do obliczeń rozproszonych z liczbą 21 rdzeni, skraca czas złożonych obliczeń 15-krotnie. W zależności od liczby przeprowadzanych analiz numerycznych, rosnący liniowo i dość szybko czas w przypadku obliczeń na pojedynczej jednostce, wykazuje niemalże stałą i stosunkowo niedużą wartość dla analogicznych obliczeń testowych z wykorzystaniem 22 rdzeni. Obserwuje się przy tym, spodziewany i potwierdzający ekonomiczne zalety obliczeń rozproszonych, spadek tzw. narzutu wraz ze rosnącym stopniem skomplikowania problemu obliczeniowego (np. liczbą analizowanych przypadków). Wszystkie powyższe trzy zjawiska, tj. znaczące skrócenie czasu obliczeń, praktyczna niezależność od liczby przeprowadzonych symulacji oraz przyzwoity narzut, zaobserwowane również dla alternatywnego środowiska z architekturą *HTCondor*, utwierdziły Autora dysertacji w przekonaniu o celowości zastosowania chmury obliczeniowej do wymagających dużego nakładu obliczeniowego licznych symulacji dla danych o charakterze stochastycznym.

W rozdziale czwartym opisane zostały trzy eksperymenty numeryczne, przeprowadzone na zasymulowanych obiektach biologicznych, w których Autor dysertacji wykorzystał zaimplementowany przez siebie system obliczeniowy oparty o chmurę publiczną, oferowaną przez firmę Microsoft (*Microsoft Azure*). W pierwszym przykładzie analizowano grzanie rezystancyjne około 100 jednorodnych obiektów biologicznych (kostka ziemniaczana, różne gatunki

ziemniaków). W przeprowadzonym eksperymencie charakter stochastyczny miały: liczba kostek i jej wymiary geometryczne, jej losowe ułożenie w pojemniku wypełnionym cieczą przewodzącą, kąt obrotu oraz konduktywność. Ta ostatnia była przedmiotem pomiarów i analiz, w celu uzyskania danych o charakterze statystycznym (np. histogramy wartości konduktywności), analogicznych do opisanych w rozdziale drugim. Kolejne dwa przykłady dotyczą elektrostymulacji nerwowo-mięśniowej (EMS), gdzie badano dwuwymiarowy model przekroju ramienia oraz terapii elektrowstrząsowej (ECT), gdzie badano trójwymiarowy model głowy. Dane geometryczne do stworzenia modeli cyfrowych, do ostatnich dwu analiz oraz niezbędne parametry „materiałowe” pobrano z literatury oraz baz ogólnodostępnych (np. WHO). Doktorant wykorzystał tu zdobyte wcześniej doświadczenia w ramach własnej akwizycji danych, by uwzględnić stochastyczny charakter danych, przy niestety nielicznych przypadkach klinicznych zaczerpniętych z baz WHO. Przykłady obliczeniowe jednoznacznie potwierdziły zasadność prowadzenia obliczeń rozproszonych i ich dużą skuteczność, wobec obliczeń na pojedynczej maszynie (1 rdzeń), mierzoną znacznie mniejszym czasem obliczeń, obniżającym się narzutem i ostatecznie także akceptowalnymi kosztami (w USD) korzystania z chmury publicznej (w tym przypadku *Microsoft Azure*). Za takim rozwiązaniem, obok wspomnianych już zalet, przemawia ogólna cecha obliczeń rozproszonych w chmurze, zapewniająca skalowalność i elastyczność, dla zadań wymagających zmiennych mocy obliczeniowych.

Symulacji komputerowych omówionych w rozdziale czwartym nie można właściwie ocenić i docenić bez zapoznania się z treścią załącznika 2 („Wdrożenie w *Microsoft Azure*”). W mojej opinii załącznik ten winien być potraktowany jako pełnoprawny, kolejny, istotny rozdział merytoryczny i znaleźć się w pracy tuż za rozdziałem trzecim. Ten zawierający osobiste, praktyczne doświadczenia Doktoranta załącznik, zawiera kluczowe dla pracy informacje o konfiguracji narzędzia obliczeniowego, wykorzystującego wirtualizację w chmurze *Microsoft Azure*. Autor rozprawy wniósł istotny wkład pracy koncepcyjnej i technicznej w celu odpowiedniego, optymalnego skonfigurowania zestawu węzłów (*Master Node/Worker Node*), umieszczenia obrazu systemu operacyjnego oraz zainstalowania niezbędnego oprogramowania symulacyjnego, wykorzystującego m.in. metodę elementów skończonych (MES).

Zbyt lakoniczny, w mojej opinii, rozdział piąty zawiera treści podsumowujące i wnioski, w ramach których Doktorant wskazuje na te elementy swojej pracy, które świadczą o słuszności sformułowanych na początku recenzowanej dysertacji założeń oraz o zrealizowaniu celu badawczego. Autor dostrzega również problemy otwarte związane z tematyką pracy i daje wskazówki dotyczące ewentualnych przyszłych badań. Ten ostatni, bardzo pożądaný element jest niestety nieczęsto spotykany w pracach doktorskich.

Pozycje bibliograficzne, drukowane zostały dobrane właściwie. Częścią bibliografii są linki do stron internetowych. Wobec oferowanych coraz nowszych narzędzi i usług informatycznych, elektronicznych serwisów publikacji naukowych, a także wymogów współczesnego marketingu, który realizowany jest głównie za pomocą Internetu, takie rozwiązanie jest w pełni uzasadnione.

Układ logiczny pracy oraz proporcje jej poszczególnych rozdziałów jest typowy dla rozpraw doktorskich. Praca, mimo pewnych, zastosowanych skrótów myślowych nie wymaga uzupełnień, ani zmian. Jej wartość merytoryczną oceniam pozytywnie. Opierając się na załączonych w pracy wynikach symulacyjnych i wynikach eksperymentów pomiarowych można uznać, że zasadniczy cel badawczy i cele pośrednie zostały osiągnięte, przy czym Autor wykazał się umiejętnością samodzielnego rozwiązywania problemów naukowych oraz szeroką wiedzą w obszarze modelowania wybranych zagadnień bioelektromagnetyzmu w ujęciu stochastycznym, z użyciem chmury obliczeniowej, możliwych do wykorzystania w szeroko rozumianej informatyce, a także w innych dyscypli-

nach, w których niezbędne jest wykorzystanie wirtualizacji w chmurze, celem uzyskania dużej mocy obliczeniowej.

3. Uwagi ogólne i szczegółowe

Czytając pracę mgr inż. Artura Sławomira Krupy, nasunęło mi się kilka wątpliwości natury merytorycznej, które pozwolę sobie przytoczyć:

Na stronie 27, 2d, czytamy:

Powyższe równanie zawiera w sobie znane nam już wartości, a są to: powierzchnia styku (s) oraz wymiar odległości między elektrodami, czyli wysokość próbki (l). Wiedząc, że każda próbka będzie w formie prostopadłościanu, wymiar l będzie jednocześnie wysokością próbki, jak i po jej obroceniu – podstawą s . Możemy zatem potroić każdy pomiar uzyskując trzy razy więcej wyników pomiarowych ze wszystkich próbek.

Jest to daleko idący i nie jedyny w recenzowanej pracy, skrót myślowy. Rysunek 13 i zależność (14) wyraźnie potwierdzają, że wymiar l jest liniowy [m], podstawa s zaś jest mianowana w jednostce odpowiadającej polu powierzchni [m²].

Rysunek 19. O ile histogram po lewej stronie nie budzi zastrzeżeń (kostka mięsno-tłuszczowa jest większa, zatem jest ona prawdopodobnie niejednorodna), to wyjaśnienia wymaga histogram po stronie prawej (kostka jest mniejsza, zapewne bardziej jednorodna). Prawy histogram, mimo pewnego uśrednienia, wyraźnie ujawnia dwa ekstrema: ok. 0,025 [S/m] oraz ok. 0,09 [s/m]. Jak to wytłumaczyć? Jeśli kostki były bardziej jednorodne, co jest logiczne (8x mniejsza objętość), to być może nastąpiła swoista „separacja” i część z nich reprezentowała zdecydowaną przewagę tkanki tłuszczowej, inna część - tkanki mięsnej?

Str. 58, 1g. Pada stwierdzenie:

Klaster VMSS zbudowany w Microsoft Azure został przygotowany z 40 maszyn, każda z 8 rdzeniami CPU, co łącznie dało 320 węzłów roboczych (Workerów). Teoretyczne przyspieszenie w takiej sytuacji powinno wynieść 320 razy. Pojedyncza symulacja zajmuje w przybliżeniu 5 sekund. Dla 100 000 przypadków zając to powinno 500 000 sekund (czyli 139 godzin). HTCondor w powyższej konfiguracji rozdystrybuował zadanie i wykonał je w czasie około 2 godzin, co pozwala wywnioskować, że narzut sięgał w przybliżeniu 150%.

Czy wszystkie powyższe dane są poprawne? Czy rzeczywiście jest to 150%?

Str. 62, 4g. Pada stwierdzenie, dotyczące prawa Joule’a i zależności (19):

Z prawa tego wynika wprost zasada zachowania energii, która mówi, że straty energii przepływającego prądu zamienia się w energię wewnętrzną.

Niestety, to zdanie nie jest prawdziwe. To prawo Joule’a jest konsekwencją zasady zachowania energii - nie odwrotnie (!)

Str. 79, 11d. Pada stwierdzenie:

Badania przeprowadzone w ramach tej pracy pokazały, że zaprezentowane podejście stochastyczne prowadzi do dokładniejszych wyników, niż wykorzystanie klasycznych modeli.

Z pewnością modelowanie stochastyczne powinno dawać wyniki bardziej adekwatne w porównaniu z wynikami uzyskanymi dzięki modelowaniu klasycznemu (deterministycznemu?). W pracy nie mogę się jednak doszukać analizy porównawczej w tym zakresie, czyli takiej,

która zestawia błąd obliczeń dla danego modelu stochastycznego (przy różnej, rosnącej liczbie przeliczonych przypadków) oraz deterministycznego (pojedyncze obliczenia autorskie lub zaczerpnięte z literatury), z wynikami, które Autor rozprawy może uznać się za referencyjne (eksperyment pomiarowy, zaczerpnięty z literatury wiarygodny eksperyment numeryczny, inne?).

Niektóre z powyższych uwag i komentarzy mają zapewne charakter dyskusyjny. Będę wdzięczny Doktorantowi za wcześniejsze (przed publiczną obroną), pisemne rozwianie moich wątpliwości.

4. Uwagi redakcyjne

Szata graficzna rozprawy doktorskiej mgr inż. Artura Sławomira Krupy jest, z małymi wyjątkami wykonana dość starannie. Szkoda, że czytelność niektórych rysunków, nawet w drukowanej wersji A4 nie może być uznana za zadowalającą. Konieczne było powiększenie dokumentu, który udostępniono mi również w wersji elektronicznej.

Praca zawiera obszerną, reprezentatywną i cytowaną literaturę, cel badawczy, opis realizacji celów szczegółowych zawarty w rozdziałach merytorycznych oraz syntetyczne podsumowanie.

Niestety, lektury pracy nie można uznać za sprawę łatwą. Rozdziały pracy przypominają bardziej lakoniczne sprawozdania z etapów wykonywanych projektów badawczych, niż ciąg myślowy, charakterystyczny dla typowej dysertacji, w której jej autor „rozprawia się” z problemem naukowym, a kolejność poszczególnych rozdziałów i ich treści jest konsekwencją pewnej myśli wiodącej. Można to próbować wytłumaczyć charakterem rozprawy przygotowanej w dyscyplinie naukowej „Informatyka” (techniczna), gdzie zależności matematyczne oraz szczegółowe opisy uzyskanych wyników, często ustępują miejsca stosunkowo zwięzłym, a może nawet ascetycznym zapisom i podsumowaniom. Doszukuję się tu pewnej analogii do formalnych, ascetycznych zapisów kodu źródłowego w programach komputerowych, którymi chętnie posługują się programiści ☺.

Autor rozprawy nie ustrzegł się drobnych błędów redakcyjnych, w szczególności tzw. „literówek”. Za mankament uznaję pewne niedoskonałości natury językowej (składniowej), występujące głównie w części wstępnej recenzowanej dysertacji.

Zwracam też uwagę na inne, drobne mankamenty:

Ogólnie. Nieuzasadnione, powszechne używanie terminu „wyliczenia” zamiast np. „obliczenia”; czy „wyliczono” zamiast „obliczono” lub „wyznaczono”. „Wyliczanie” jest w moim przekonaniu pojęciem odpowiadającym takim czynnościom jak „wymienianie” lub „wypunktowywanie” (czegoś).

Ogólnie. Podany opis „technologii” skonfigurowania maszyny wirtualnej w chmurze *Microsoft Azure*, może się okazać bardzo pomocny dla ewentualnych „naśladowców”. Jednak z uwagi brak danych nie jest możliwe powtórzenie obliczeń w tej samej, czy też innej, analogicznej chmurze obliczeniowej.

Rysunek 3. Nie jest dla mnie jasne, czemu służy prezentacja tego rysunku?

Rysunek 57. Czy ten „rysunek” nie powinien być zaklasyfikowany jako tabela?

Rysunki. Wątpliwości budzą anglojęzyczne teksty w rysunkach, polskojęzycznej pracy. Choć większość z tych rysunków została zaczerpnięta z anglojęzycznych źródeł (stron internetowych), za zasadne wydaje się znalezienie (o ile istnieją?) i podanie polskojęzycznych odpowiedników.



Na przykład rysunki. Niektóre z rysunków Autor zaczerpnął z Internetu, podając w podpisie ich źródło (link). O ile jest to uzasadnione, np. w przypadku internetowych stron usługodawców, oferujących dostęp do chmur obliczeniowych, to już podawanie jako wiarygodnego źródła wiedzy stron z serwisu Wikipedia wydaje się być ryzykowne. W pracach naukowych należy raczej sięgać do sprawdzonych źródeł oryginalnych (czasopisma, doniesienia konferencyjne, normy, instrukcje techniczne etc.) lub monografii naukowych, niż do informacji wtórnych i znacznie przetworzonych, jak te zawarte w Wikipedii, czy np. w skryptach dla studentów.

Literatura. Nie wszystkie dane bibliograficzne są kompletne. W szczególności przy niektórych pozycjach brakuje stron publikacji. Być może wynika to z faktu, iż niektóre materiały konferencyjne są udostępniane jedynie na płycie CD i poszczególne artykuły są umieszczane w odrębnych plikach (?).

5. Wniosek końcowy

W mojej ocenie mgr inż. Artur Sławomir Krupa podjął się ciekawego i ważnego z punktu widzenia praktyki obliczeń symulacyjnych zadania, opracowania od strony teoretycznej i wdrożenia w publicznej chmurze obliczeniowej *Microsoft Azure* systemu obliczeniowego, bazującego m.in. na *solverach* wykorzystujących MES. O wartości opracowania decyduje nie tylko dokonane wdrożenie, które dzięki zrównolegleniu obliczeń umożliwia rozwiązywanie licznych zadań wymagających odpowiednio dużych mocy obliczeniowych, lecz także opracowanie techniki obliczeń uwzględniających aspekty stochastyczne. Typowe symulacje MES ograniczają się najczęściej do obliczeń deterministycznych. Sprawdza się to w stosunku do projektowania urządzeń, dla których parametry materiałowe lub ich ewentualna zmienność w czasie, czy też zmienność wynikająca z nieliniowości są znane lub bardzo dobrze przewidywalne. Tymczasem w wielu wypadkach, np. w odniesieniu do badań struktur biologicznych, a w szczególności badań w ramach bioelektromagnetyzmu, konieczne jest uwzględnienie stochastycznego charakteru zmian parametrów „materiałowych”, wymiarów geometrycznych, położenia w przestrzeni i innych.

Do głównych osiągnięć Doktoranta, należą autorskie opracowanie i zaimplementowanie:

- metody odtwarzania funkcji gęstości prawdopodobieństwa parametrów materiałowych tkanek na podstawie niewielkiej liczby danych pomiarowych oraz przygotowanie programów analizy statystycznej, których celem było odtwarzanie funkcji gęstości prawdopodobieństwa, na podstawie niewielkiej liczby danych pomiarowych,
- umiejętne przeniesienie doświadczeń zdobytych w ramach akwizycji danych „materiałowych”, tkanek mięsno-tłuszczowych pochodzenia zwierzęcego oraz z utworzenia własnej, licznej bazy danych o charakterze stochastycznym, na rzeczywiste, lecz ograniczone liczebnie dane z serwisu WHO, dotyczące tkanek ludzkich,
- metody porównywania wydajności rozproszonych systemów obliczeniowych w różnych konfiguracjach,
- skalowalnego systemu informatycznego do obliczeń rozproszonych, zbudowanego w oparciu o autorskie obrazy systemów operacyjnych,
- programów symulacyjnych FEM, przygotowanych do obsługi parametrów, opisywanych zmiennymi losowymi, działających w chmurze, w oparciu o środowisko *FEniCS*,

a ponadto umiejętne utworzenie adekwatnych modeli matematycznych oraz modeli cyfrowych do stochastycznej analizy, w ramach trzech przykładów obliczeniowych z zakresu bioe-

lektromagnetyzmu, których wyniki potwierdzają metodologiczną i ekonomiczną zasadność wykonywania obliczeń rozproszonych w chmurze.

Powyższe elementy, wśród których najważniejsze i najbardziej pracochłonne są ściśle powiązane z dyscypliną naukową „Informatyka”, były niezbędne do zrealizowania zasadniczego celu badawczego.

Uwagę zwracają duże umiejętności praktycznego zastosowania przez Doktoranta rozważań teoretycznych, których dotyczy dysertacja, w rozwiązaniach opartych na obliczeniach w chmurze, w szczególności utylitarne elementy wykonanych prac, powiązane z szeroko rozumianą informatyką stosowaną:

- demonstracja możliwości nietypowego wykorzystania platformy *Apache Hadoop* do dystrybucji zadań obliczeniowych,
- opracowanie wyspecjalizowanej, dedykowanej do konkretnych zastosowań konfiguracji chmury obliczeniowej *Microsoft Azure*,
- demonstracja wykorzystania komercyjnego zaplecza serwerowego, oferowanego jako usługi chmury obliczeniowej do obliczeń naukowych z wykorzystaniem platformy *Microsoft Azure*,
- decydująca o zaletach ekonomicznych przyjętego rozwiązania, analiza tzw. narzutów.

Należy ponadto docenić:

- trafne dostrzeżenie przez Doktoranta problemów jeszcze nie rozwiązanych, świadczących o jego szerokiej wiedzy oraz wskazujących na potrzebę kontynuacji badań i pokazujących drogę naukową następcom,
- dobre opanowanie warsztatu naukowego.

Wydaje się, że praca może zainspirować specjalistów z innych obszarów nauki, np. wykonujących obliczenia polowe, wymagające danych o charakterze stochastycznym.

Autor wykazał się dobrą wiedzą teoretyczną i rozeznaniem literaturowym w zakresie interdyscyplinarnym (informatyka/elektrotechnika/bioinżynieria) oraz szerokiego zastosowania technik modelowania komputerowego w chmurze obliczeniowej, w tym możliwych do wykorzystania w zagadnieniach bioelektromagnetyzmu. W recenzowanej pracy, której fragmenty były opublikowane w liczących się czasopismach naukowych, dostrzegam wiele oryginalnych osiągnięć jej Autora, świadczących o wystarczających kompetencjach Doktoranta do samodzielnego rozwiązywania problemów naukowych oraz wskazujących na jego wiedzę ogólną i umiejętności praktyczne, niezbędne w dyscyplinie naukowej „Informatyka”.

Konkluzja:

Powyższe uwagi upoważniają mnie do stwierdzenia, że praca pt. „Metody obliczeń rozproszonych w stochastycznym modelowaniu tkanek” autorstwa mgr inż. Artura Sławomira Krupy, złożona w przewodzie doktorskim w dyscyplinie naukowej „Informatyka”, spełnia wymogi, stawiane rozprawom doktorskim, zgodnie z Ustawą z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2003r. Nr 65, poz. 595, z późniejszymi zmianami) i stawiam wniosek o dopuszczenie jej do publicznej obrony.

Szczecin, dnia 23 kwietnia 2019 r.