

Poznań, 20.06.2022 r.

Dr hab. inż. Krzysztof Kurowski
Instytut Chemii Bioorganicznej PAN
Poznańskie Centrum Superkomputerowo – Sieciowe
Ul. Jana Pawła II 10
61-139 Poznań

Akademia Górniczo – Hutnicza
im. Stanisława Staszica w Krakowie
Rada Dyscypliny Informatyka
Techniczna i Telekomunikacja
Al. A. Mickiewicza 30
30-059 Kraków

Recenzja

Przedmiotem recenzji jest rozprawa mgra inż. Mateusza Paciorka zatytułowana w języku angielskim „*Scalable simulation of social phenomena based on signal propagation modeling*”. Promotorem rozprawy jest dr hab. inż. Wojciech Turek. Recenzja została opracowana na podstawie zlecenia RD.ITiT.WIEiT.510-25/19/117/2022 działającego na podstawie uchwały Rady Dyscypliny Informatyki Technicznej i Telekomunikacji AGH z dnia 20.04.2022 r.

1. Znaczenie podjętej tematyki oraz cel rozprawy

Modelowanie oparte na agentach jest ważną techniką modelowania symulacyjnego, która pozwala zbadać, w jaki sposób interakcje heterogenicznych jednostek wpływają na szersze zachowanie systemów społecznych i czasowo-przestrzennych. Warto podkreślić fakt, iż dzięki rozwojowi systemów obliczeniowych oraz sztucznej inteligencji, technika modelowania i symulacji oparta na agentach nadal zyskuje szerokie grono zwolenników oraz rosnącą liczbę praktycznych zastosowań, zaczynając od symulacji interakcji mikroorganizmów, aż po złożone, dynamiczne i często długotrwałe procesy oraz zjawiska obserwowane w naturze, ze szczególnym uwzględnieniem i zrozumieniem procesów społecznych oraz rozproszonych procesów decyzyjnych człowieka na poziomie jednostki, grupy, czy nawet całych złożonych populacji. Badania symulacyjne w oparciu o różne agentowe podejścia dostarczają najczęściej podstawowej wiedzy na temat interakcji różnych jednostek w złożonych systemach

modelowanego świata rzeczywistego. Dodatkowo, modelowanie i symulacje agentowe umożliwiają nie tylko ocenę działania samego systemu, ale w stosunkowo łatwy sposób pozwalają na uwzględnienie i pogłębioną analizę różnych parametrów kontrolnych, czynników i ograniczeń mających wpływ na zachowanie się w czasie i przestrzeni badanego systemu oraz wzajemną interakcje i oddziaływanie wszystkich jednostek składowych. Pozyskana w ten sposób wiedza badanego systemu pozwala na wykorzystanie symulacji agentowych w wielu ciekawych zastosowaniach, w tym może być z powodzeniem wykorzystana do rozwiązywania problemów z dziedziny nauk inżynierijno-technicznych (np. architektury, urbanistyki, transportu i inżynierii lądowej), nauk społecznych (np. ekonomii, gospodarce przestrzennej, bezpieczeństwie) oraz nauk ścisłych i przyrodniczych (np. informatyce i biologii). Tym samym, należy zaznaczyć, iż prezentowana tematyka badawcza związana z modelowaniem i symulacjami w oparciu o agentowe podejście, w szczególności w ich praktycznym ujęciu, wymaga silnie interdyscyplinarnego ujęcia dla analizowanego problemu. Jednocześnie, rozważana problematyka wymaga od użytkownika znajomości wielu zagadnień z dziedziny samej informatyki, np. algorytmiki, inżynierii oprogramowania, architektury komputerów i systemów superkomputerowych, a w kontekście stawianych wyzwań naukowych w pracy doktorskiej, również dobrej znajomości zasad programowania równoległego i rozproszonego. Podjęta przez Doktoranta próba stworzenia wysoko skalowalnej, a jednocześnie abstrakcyjnej metody dla reprezentacji dyskretnych i bardziej złożonych przestrzennych modeli w oparciu o agentowe podejścia, przy jednoczesnym kluczowym założeniu zachowania poprawności oraz powtarzalności wykonywanych symulacji na systemach komputerowych dużej mocy (superkomputerach), niezależnie od rozmiaru instancji, nie jest w istocie zadaniem trywialnym. Główna teza doktoratu mgra inż. Mateusza Paciorka została sformułowana poprawnie i w przemyślany sposób akcentując stosownie trzy główne wyzwania badawcze, odpowiednio:

- zapewnienie, że zaproponowana metoda nie będzie miała wpływu na model rozproszony i zmiany jego stanu, czyli zachowana zostanie jego poprawność niezależnie od stopnia zrównoleglenia symulacji agentowych na systemie superkomputerowym;
- zachowana będzie wysoka skalowalność metody uruchamianej w infrastrukturze superkomputera;
- metoda ułatwi modelowanie bardziej złożonych struktur agentowych (bez typowych ograniczeń wynikających z podstawowej reprezentacji w modelach agentowych opartej na dwuwymiarowej siatce), a jednocześnie pozwoli na zwiększenie potencjalnych scenariuszy praktycznego wykorzystania podejścia modelowania i symulacji agentowych.

Doktorant we wstępnych rozdziałach rozprawy w sposób merytoryczny charakteryzuje główne założenia oraz podejścia do modelowania i symulacji agentowych. Następnie, umiejętnie prezentuje ich liczne ograniczenia, akcentując kluczowe dla podejmowanych przez niego badań naukowych ograniczenie występujące w wielu środowiskach agentowych opartych o dwuwymiarową siatkę z relatywnie statycznymi strukturami danych i algorytmami nie wspierającymi zrównolegionych i skalowalnych symulacji komputerowych. Na tej podstawie, przy wykorzystaniu stosownego aparatu matematycznego, Doktorant wprowadza dodatkowe definicje oraz proponuje stosowne usprawnienia doprowadzając w efekcie swoich rozważań do poprawnego opisu podstawowych założeń dla nowej, skalowalnej metody modelowania środowiska agentowego przy zachowaniu jej pełnej kompatybilności z referencyjnymi uproszczonymi podejściami. W kolejnym etapie swoich rozważań, Doktorant poprawnie przedstawia bazowy schemat testowania poprawności zaproponowanej metody. W tym celu dobiera on znane w przedmiotowej literaturze metody testowe wykorzystywane do oceny skalowalności i poprawności dla rozważanych różnych konfiguracji eksperymentów obliczeniowych uwzględniających zarówno zwiększającą się liczbę elementów/rdzeni

przetwarzania w infrastrukturze superkomputera, jak i zwiększający się rozmiar instancji wybranych problemów testowych. Warto zaznaczyć, iż Doktorant również w przejrzysty i czytelny sposób przedstawia przykładowe modele agentowe oparte na rzeczywistych scenariuszach wykorzystujących zaproponowaną metodę, a także omawia dodatkowe rozszerzenia związane z dalszym poszerzaniem modeli agentowych, które mogą korzystać z proponowanej metody.

Reasumując, podjęty w rozprawie doktorskiej zakres prac naukowych, zarówno od strony teoretycznej, jak i praktycznej, wpisuje się w silny nurt naukowo-badawczy w dyscyplinie Informatyka związany z modelowaniem i symulacjami agentowymi. Podjęta przez Doktoranta tematyka wykorzystania rosnącego potencjału i mocy systemów superkomputerowych w powiązaniu z opracowaniem nowej, skalowalnej, poprawnej i bardziej elastycznej metody do modelowania i symulacji agentowych jest ważna z perspektywy naukowej. Zaprezentowane wyniki prac badawczych Doktoranta stanowią istotny oraz twórczy wkład w rozwój skalowalnych modeli i symulacji komputerowych w oparciu o agentowe podejścia, czego dowodem jest również cykl publikacji Doktoranta, do których znajdują się liczne odniesienia w rozprawie doktorskiej. Zaprezentowany zakres działań naukowych stopniem trudności oraz oryginalności odpowiada w pełni wymaganiom, jakie zwyczajowo i ustawowo stawia się rozprawom doktorskim w dziedzinie Informatyka.

2. Analiza i ocena poszczególnych fragmentów pracy

Rozprawa doktorska mgra inż. Mateusza Paciorka to spójny tematycznie opis oryginalnej pracy naukowej bazującej na opublikowanych w latach 2018 – 2021 dziewięciu artykułach naukowych, w tym trzech opublikowanych w uznanych czasopismach (punktacja MNiSW \geq 100 pkt). Przedłożona rozprawa doktorska składa się z 9 tematycznych Rozdziałów głównych oraz załączników opisujących szczegółowe scenariusze testowe wykorzystane w eksperymentach obliczeniowych do weryfikacji i oceny skalowalności oraz poprawności zaproponowanej metody (łącznie 100 stron wraz z bibliografią). W samej rozprawie można wyróżnić zasadniczo dwie części, odpowiednio Rozdziały 1-4 skupiające się na wprowadzeniu podstawowych założeń i zagadnień wraz z przykładowymi ilustracjami związanymi z modelowaniem i symulacjami w oparciu o agentowe podejścia oraz stosownymi uzasadnieniami potrzeby rozbudowy i usprawnień istniejących podejść. Następnie, przedstawione Rozdziały 5-8 odnoszą się już bezpośrednio do głównej tezy rozprawy doktorskiej oraz powiązanych z nią obszarów badań teoretycznych i eksperymentalnych. Rozdział 9 to zwięzłe i merytoryczne podsumowanie osiągniętych wyników pracy naukowej Doktoranta wraz z istotnymi kierunkami dalszych prac badawczych.

Główne motywacje oraz założenia rozprawy doktorskiej, w tym kluczowa teza naukowa oraz najistotniejsze cele dla podjętych badań naukowych wraz z krótkim podsumowaniem oryginalnego wkładu Doktoranta zostały przedstawione w Rozdziale 1. W tym wprowadzającym rozdziale lub ew. w Rozdziale 2 niestety zabrakło pogłębionego przeglądu literaturowego w odniesieniu do różnych typów i konfiguracji systemów superkomputerowych, w szczególności odniesienia się przez Doktoranta do wyzwań związanych z rozproszoną komunikacją i synchronizacją wiedzy pomiędzy agentami (np. wyzwania związane z „...the influence of delayed synchronization and data loss in communication, inter-process synchronization, buffer zones and signal propagation”, itd.) na tle różnych topologii połączeń sieciowo-komunikacyjnych i związanych z tym opóźnień

komunikacyjnych, heterogenicznych i hierarchicznych układów przetwarzania, akceleratorów sprzętowo-programowych, itp. Taka wstępna analiza jest istotna z perspektywy rozważanej problematyki skalowalności, oceny stopnia zrównoleglenia nowej metody oraz doboru odpowiedniego agentowego środowiska symulacyjnego. W efekcie takich rozważań, łatwiej można byłoby zrozumieć przyjęte przez Doktoranta założenia związane z weryfikacją skalowalności, eksperymentami obliczeniowymi oraz ostatecznie wyborem tylko jednego typu systemu superkomputerowego i środowiska agentowego. Wskazany system superkomputerowy o ściśle określonej topologii połączeń sieciowo-komunikacyjnych i charakterystyce wydajnościowo-energetycznej układów przetwarzania wraz z wybranym środowiskiem agentowym spełnia podstawowe wymagania stawiane w rozprawie doktorskiej, ale nie pokrywa pełnego spektrum możliwych konfiguracji sprzętowo-programowych dla zrównolegionych i rozproszonych agentowych środowisk symulacyjnych uruchamianych w infrastrukturze superkomputerów. Jest to o tyle istotne, iż pozwoliłoby to na przedstawienie wyników pracy naukowej Doktoranta w znacznie szerszym kontekście badawczym, a jednocześnie zwiększyłoby potencjalnie zakres zastosowań opracowanej metody („... *the method must not limit its applicability to a narrow class of ABMS, instead providing as much flexibility and extensibility as possible*”). Doktorant wprowadza w rozdziale 5.3. *Distribution and efficiency of the proposed method* dokonuje analizy związanej z poprawą efektywności komunikacji pomiędzy agentami, optymalizacji struktur i synchronizacji pomiędzy różnymi fazami symulacji oraz zwiększania poziomu zrównoleglenia modeli, ale rozważania odnoszą się tylko do warstwy aplikacyjnej w oderwaniu od warstwy sprzętowo-programowej wspieranej w różnych konfiguracjach przez systemy superkomputerowe.

Relatywnie bardziej obszerny przegląd literaturowy został przedstawiony w odniesieniu do istniejących podejść oraz narzędzi wspierających modelowanie i symulacje oparte na agentowych technikach w Rozdziale 2. Niestety, również i w tym przypadku, Doktorant bardzo mocno ograniczył swoje rozważania i analizy funkcjonalne istniejących agentowych środowisk symulacyjnych wskazując tylko na wybraną bazową platformę Xinuk (środowisko oparte o wirtualne maszyny Javy) oraz wymieniając dwa agentowe środowiska symulacyjne FLAME i Repast z pośród wielu dostępnych i opisanych w literaturze agentowych środowisk symulacyjnych.

Rozdział 3 to zbiór wstępnych rozważań oraz kierunkowych prac badawczych podjętych przez Doktoranta zaprezentowany w publikacjach naukowych w celu zrozumienia i opracowania właściwej definicji poprawności metody symulacyjnej. W tym kontekście warto byłoby pogłębić rozważania o referencyjne techniki kwantyfikacji niepewności oraz projektowania eksperymentów symulacyjnych, które mogą pomóc jeszcze lepiej zrozumieć złożoną naturę systemów agentowych w modelach obliczeniowych. Takie analizy pozwalają odkryć nowe zależności między danymi wejściowymi i wyjściowymi badanego modelu, uwzględnić różne rodzaje i tryby projektów eksperymentalnych, oceny niepewności i wrażliwości modelu, która pomaga zidentyfikować części modelu mające znaczenie w eksperymentach, a także narzędzia statystyczne do kalibracji modeli w odniesieniu do dostępnych danych.

Rozdział 4 odnosi się do istotnego zagadnienia rozważanego przez Doktoranta w swojej pracy naukowej związanego z bardziej elastyczną reprezentacją struktur danych w środowisku agentowym i propozycji przejścia z reprezentacji struktury siatki 2D, na bardziej ogólną strukturę grafów zaadaptowanej do nowej metody propagacji sygnałów. Główną zaletą zaproponowanego ulepszenia jest odejście od obliczeń opartych na wskaźnikach, podatnych na błędy, na rzecz bardziej intuicyjnych, opartych na kierunkach. Jednocześnie Doktorant słusznie podkreśla zalety nowego i bardziej elastycznego podejścia wskazując na jego istotne znaczenie w zakresie możliwości modelowania agentów w przestrzeni 3D, typowego założenia dla większości scenariuszy świata i systemów rzeczywistych.

W Rozdziale 5 zaprezentowano szczegółowy opis i definicję skalowalnego modelu, a następnie przedstawiono algorytm aktualizacji i synchronizacji wykorzystywany w proponowanej metodzie wraz z ilustracją i przykładem. Doktorant podzielił się wieloma ciekawymi obserwacjami wskazującymi nie tylko na zdobyte doświadczenie w zakresie praktycznego wykorzystania modelowania i symulacji agentowych, ale w przemyślanym przedstawianiu technik zwiększania poziomu ich skalowalności. Wskazane praktyki obejmują zarówno zasady niezmienności skali i utrzymywania stałej liczby procesów, regularny podział środowiska, odpowiedni dobór rozmiaru granic regionów i związanych z nimi procesów komunikacji i synchronizacji symulacji, itp. Niemniej jednak, jak zostało to już wspomniane, w przypadku zwiększającej się heterogeniczności i hierarchiczności systemów superkomputerowych oraz węzłów komunikacyjnych. Bardziej złożone metody podziału środowiska uwzględniające topologie i charakterystyki sprzętowo-programowe superkomputerów oraz metody zarządzania dynamicznym poziomem obciążenia obliczeniami na różnych układach przetwarzania mogą zapewnić jeszcze lepszy poziom skalowalności i wydajności (w tym wydajności energetycznej, nie tylko czasu trwania symulacji) dla rozważanych środowisk symulacji agentowych. Doktorant ograniczania tego typu rozważania tylko do wskazania ich jako istotnych, ale w dalszych pracach naukowych.

Schemat weryfikacji i wyniki eksperymentów obliczeniowych zostały zaprezentowane w Rozdziale 6 w poprawnym schemacie walidacyjnym: analiza modelowanego scenariusza i wybór kluczowych metryk opisujących najważniejsze obserwowane zachowania; wielokrotne wykonanie obliczeń i odczyt wybranych metryk z wielu powtórzeń o różnym stopniu rozłożenia, przy zachowaniu tej samej konfiguracji parametrów oraz analiza statystyczna zebranych danych, w szczególności dla podejść i modeli niedeterministycznych. Doktorant zaprezentował również w formie ilustracji graficznej i nagrania wyniki działania metody w eksperymencie obliczeniowym dla przykładowego modelu deterministycznego uruchomionego w wersji na 1 rdzeniu i w wersji zrównoleglonej na 100 rdzeniach, wskazując na pełną poprawność pomiędzy uzyskanymi wynikami. O ile wizualizacja procesu symulacji agentowej jest niezbędnym elementem w analizie wyników, to w przypadku porównania różnych uruchomień warto byłoby też zaprezentować jakąś prostą i zagregowaną miarę ilościową dla dwóch lub większej liczby eksperymentów obliczeniowych pokazującą stopień poprawności na etapach realizacji symulacji.

Rozdział 7 przedstawia koncepcyjnie spójnie i metodologicznie poprawne eksperymenty obliczeniowe w zakresie badania skalowalności zaproponowanej metody z uwzględnieniem przyjętych w literaturze metod pomiaru skalowalności kodów aplikacyjnych, w tym przyspieszenia obliczeń w eksperymentach silnej i słabej skalowalności oraz metryki Karp-Flatta. Zakres eksperymentów Doktorant ograniczył tylko do jednej infrastruktury superkomputera w określonej architekturze i topologii połączeń.

W kolejnym Rozdziale 8 znajdują się obszernie i bardzo ciekawe opisy przykładów zastosowania zaproponowanej metody dla różnych rzeczywistych scenariuszy modelowania i symulacji w oparciu o agentowe techniki, w szczególności modelowaniu stosunkowo złożonych zachowań społecznych na przykładzie ruchu oraz interakcji pieszych. Ponadto, Doktorant dowodzi, iż zaproponowana metoda symulacji agentów w przestrzeni ciągłej z geometrycznym podejściem do zderzeń pozwala na występowanie zjawisk niemożliwych do uzyskania w siatce dyskretnej, a zaproponowany schemat synchronizacji i obserwacji otoczenia zapewnia dobrą skalowalność i eliminuje liczne ograniczenia, które opisywane były w rozdziałach na wstępie rozprawy doktorskiej.

Ostatni Rozdział 9 jest krótkim podsumowaniem podjętych tematów w rozprawie doktorskiej oraz wskazuje na bardzo istotne, w szczególności w zakresie wykorzystania metody w różnych środowiskach superkomputerowych dalsze kierunki pracy naukowej.

Konkludując, pomimo zwięzłej formuły opisowej nawiązującej do opublikowanych wyników prac naukowych, pewnych drobnych braków i pomyłek redakcyjno-edytorskich, zawężonego zakresu rozważań w zakresie różnych systemów superkomputerowych i środowisk symulacyjnych, **rozprawa doktorska jest merytorycznie przekonująca**. Przedstawione wyniki eksperymentów obliczeniowych dla testowych problemów potwierdzają zakładane kluczowe tezy w zakresie poprawności, skalowalności oraz możliwość zastosowania opracowanej przez Doktoranta metody w środowiskach agentowych dla bardziej skomplikowanej strukturze i większym poziomie złożoności modelowanych systemów.

3. Ocena końcowa pracy

Kandydat posiada dostateczną wiedzę w dyscyplinie Informatyka oraz nabył wymagane doświadczenie i umiejętność samodzielnego prowadzenia pracy naukowej.

Stwierdzam, że rozprawa doktorska Pana mgr inż. Mateusza Paciorka pt.: „*Scalable simulation of social phenomena based on signal propagation modeling*” **spełnia wymogi określone w Ustawie o stopniach i tytule naukowym**, w tym: art. 13 ust. 1 Ustawy z dnia 14.03.2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65 r., poz. 595 z późn. zm.).

W związku z tym **wnoszę o dopuszczenie Pana mgr inż. Mateusza Paciorka do publicznej dyskusji nad Jego rozprawą doktorską** w Radzie Dyscypliny Informatyki Technicznej i Telekomunikacji Akademii Górniczo-Hutniczej w Krakowie.

