

prof. dr hab. inż. Andrzej Obuchowicz  
Uniwersytet Zielonogórski  
Instytut Sterowania i Systemów Informatycznych

Zielona Góra, 5 stycznia 2019

## **RECENZJA**

rozprawy doktorskiej Pana mgr Oskara Wszyńskiego

*Evolutionary Algorithm for  
Particle Trajectories Reconstruction*

opracowana na zlecenie

Rady Naukowej Instytutu Podstawowych Problemów Techniki

Polskiej Akademii Nauk

### **Obszar problemowy rozprawy i istotność podjętych zadań**

Przedstawiona do recenzji praca dotyczy dwóch zastosowań algorytmu strategii ewolucyjnej z adaptacją macierzy kowariancji operacji mutacji (CMA-ES). Pierwsze z nich dotyczy rekonstrukcji trajektorii cząstek elementarnych poruszających się w polu magnetycznym zarejestrowanych w detektorze gazowym, a drugie identyfikacji modelu klasyfikatora wieloklasowego, służącego do rozpoznawania jąder powstałych w reakcjach jonów ciężkich zarejestrowanych w układzie wielodetektorowym. Oba rozważane zagadnienia są istotne z punktu widzenia fizyki cząstek elementarnych, zwanej też fizyką wysokich energii, ponieważ wiążą się z identyfikacją cząstek elementarnych rejestrowanych w detektorach akceleratora SPS (Super Proton Synchrotron). Zainteresowanie doktoranta tą tematyką jest wynikiem jego udziału w programie doktoranckim w CERN w ramach tworzenia infrastruktury dla eksperymentu NA61/SHINE.

Proponowane przez doktoranta tezy pracy są sformułowane następująco (tłumaczenie z angielskiego autora niniejszej recenzji):

„Stochastyczne algorytmy (black-box) optymalizacji pozwalają na lokalną rekonstrukcję

trajektorii cząstek elementarnych w komorze projekcyjnej w wysoką dokładnością powyżej 90%”.

oraz

„Stochastyczne algorytmy (black-box) optymalizacji pozwalają na w pełni automatyczną identyfikację cząstek powstałych w wyniku zderzenia ciężkich jonów o średniej energii, zarejestrowane przez detektory teleskopowe.”

Trudno jest określić na bazie powyższych tez, na ile podjęte zadanie jest istotne w punktu widzenia dyscypliny *Informatyka*, a na ile są one bliższe szerzej rozumianej *Fizyce instrumentalnej*.

### **Koncepcja oraz realizacja rozprawy**

Formalnie, opiniowana rozprawa doktorska, licząca 179 stron, składa się z jedenastu rozdziałów, w tym wstępu, dziewięciu rozdziałów zasadniczych i podsumowania. Do pracy dołączony jest spis rysunków i słownik występujących w pracy skrótów. Pracę uzupełnia wykaz 136 pozycji cytowanej literatury.

Pierwszy rozdział pracy stanowi stosunkowo zwięzłe wprowadzenie do pracy doktorskiej. Po krótkim przedstawieniu zagadnień rekonstrukcji trajektorii cząstek w detektorze gazowym i identyfikacji jąder atomowych, doktorant przedstawił tezy pracy i wynikające z nich zasadnicze etapy badawcze. Na końcu rozdziału znajduje się krótki opis struktury pracy i zawartości poszczególnych rozdziałów.

Rozdziały od drugiego do piątego, dotyczące kolejno ogólnego wprowadzenia do eksperymentu NA61/SHINE, wybranych urządzeń elektronicznych używanych w eksperymentach fizyki cząstek elementarnych, systemu akwizycji danych w eksperymencie NA61, oraz systemu wstępnej filtracji danych, stanowią dość obszerne nakreślenie podstawowego problemu pracy wraz z jego otoczeniem. Rozdział szósty przedstawia strukturę oprogramowania obsługującego eksperyment NA61, o nazwie SHINE, napisaną w języku C++ i obejmującą trzy główne części: zbiór modułów przetwarzania, które mogą być składane i sekwencjonowane przez użytkownika za pomocą plików XML, model danych zdarzeń, który zawiera wszystkie informacje o symulacji i rekonstrukcji oparte na strumieniu STL i ROOT oraz opis detektora, który dostarcza danych dotyczących konfiguracji i stanu eksperymentu. Aby zapewnić odpowiednią szybkość przetwarzania narzędzia, wprowadzono wrappery, które pozwalają na uruchamianie starszych części kodu pisanych dla poprzedniego eksperymentu NA49 w językach Fortran, C i innych, jako modułów w nowej strukturze.

Najistotniejszą część pracy stanowią rozdziały od siódmego do dziesiątego. W rozdziale siódmym autor przedstawił algorytm strategii ewolucyjnej z adaptacją macierzy kowariancji, który jest zasadniczym, z punktu widzenia tezy pracy, elementem procesu rekonstrukcji trajektorii cząstek. Rozdział ósmy zawiera formalne zdefiniowanie problemu rekonstrukcji trajektorii cząstek jako problemu optymalizacyjnego przedstawiając czytelnikowi metodę rekonstrukcji zdarzeń w detektorze, krótkie wprowadzenie w kinematykę cząstek i formalny model trajektorii cząstek. Rezultaty badań związanych z rekonstrukcją trajektorii cząstek zawarto w rozdziale dziewiątym. Rozdział dziesiąty dotyczy wyników badań związanych z drugim zagadnieniem podjętym w pracy, tzn. procesem identyfikacji jąder uzyskanych w wyniku zderzeń ciężkich jąder. Praca została dość lakonicznie podsumowana w rozdziale jedenastym.

Z merytorycznego punktu widzenia, biorąc pod uwagę dyscyplinę, w której ma być bronią niniejsza dysertacja, jak i postawione tezy pracy, obrona struktura pracy jest, moim zdaniem, niewłaściwa. Zwraca uwagę duża dysproporcja pomiędzy rozdziałami wprowadzającymi informacje dotyczące projektu NA61/SHINE (rozdziały 2-6, łącznie 64 strony), a rozważaniami związanymi bezpośrednio z tezą pracy, tzn. dotyczącymi stochastycznych algorytmów optymalizacji, zarówno klasycznych, heurystycznych i metaheurystycznych (rozdział 7, 9 stron). Moim zdaniem w rozdziałach 2-6 przedstawiono wiele treści nieistotnych dla podjętego problemu w pracy, zbyt szczegółowo skupiono się na technice pozyskiwania danych. Z punktu widzenia tezy pracy są to informacje nieistotne. Nabierają one znaczenia dopiero w fazie interpretacji fizycznej wyników, które w znacznym stopniu zależą od wiarygodności informacji, jaką niosą dane wejściowe, a nie mają znaczenia na proces optymalizacji ukryty w procesie rekonstrukcji trajektorii cząstek. Z drugiej strony autor postawił tezę, moim zdaniem nieuzasadnioną i niewłaściwą, że algorytm CMA-ES stanowi *state-of-the-art* w bardzo szerokiej klasie algorytmów optymalizacji stochastycznej, w szczególności algorytmów ewolucyjnych, po czym skupił się jedynie na tym algorytmie. Brak tu merytorycznej, czy eksperymentalnej, analizy porównawczej wybranego algorytmu z innymi algorytmami stochastycznymi „black-box” optymalizacji, takie jak symulowane wyżarzanie, optymalizacja rojem cząstek, czy inne z bardzo szerokiej klasy algorytmów rojowych, (pomijając już deterministyczne algorytmy bezgradientowe). Jednym z najistotniejszych wniosków płynących w twierdzenia *no free lunch* Davida Wolperta i Williama Macready’ego (1997) jest fakt, że dla zagadnienia optymalizacji globalnej nie ma uniwersalnego algorytmu poprawnego. Dlatego ocena wyższości danego algorytmu na podstawie znanych z literatury wyników badań porównawczych na pewnej klasie zadań benchmarkowych nie gwarantuje jego wyższości w rozwiązywaniu konkretnego zadania rzeczywistego. Ponadto w rozdziale podsumowującym całą pracę brak wskazania na konkretnych orygi-

nalnych osiągnięć autora w kontekście deklarowanej dyscypliny naukowej. Jest to o tyle istotne, że w pracach nad systemami wspomagającymi eksperyment NA61/SHINE bierze udział liczny zespół badaczy różnych dyscyplin naukowych. Autor w tekście pracy umieszczał krótkie notki swego wkładu, lecz warto byłoby zebrać je w konkluzjach pracy.

## **Oryginalne osiągnięcia**

Zgodnie z art. 13 ust 1 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki rozprawa doktorska, przygotowywana pod opieką promotora albo pod opieką promotora i promotora pomocniczego, powinna stanowić:

- oryginalne rozwiązanie problemu naukowego,
- lub oryginalne rozwiązanie problemu w oparciu o opracowanie projektowe, konstrukcyjne, technologiczne,
- oraz wykazywać ogólną wiedzę teoretyczną kandydata w danej dyscyplinie naukowej oraz umiejętność samodzielnego prowadzenia pracy naukowej.

W mojej opinii przedstawiona do recenzji rozprawa doktorska nie spełnia oczekiwań zawartych w powyższym artykule. W kontekście powyższego artykułu ustawy, wskazanej dyscypliny naukowej jaką jest Informatyka i przedstawionych przez doktoranta tez pracy, oczekiwałbym jednej z dwóch możliwych jej realizacji. W pierwszym podejściu winno być przedstawione oryginalne rozwiązania algorytmiczne problemu, w postaci nowatorskiego algorytmu stochastycznej optymalizacji globalnej, lub istotnej modyfikacji istniejącego już algorytmu, np. CMA-ES, pozwalającej na bardziej efektywne rozwiązanie postawionego problemu. Ewentualna wyższość proponowanego oryginalnego rozwiązania winna być zweryfikowana poprzez porównanie rezultatów z pierwotnym algorytmem, lub innymi danej klasy narzędziami. W drugim podejściu akcent nie jest postawiony na konkretne narzędzie, ale na trudny do rozwiązania problem. Oczekuje się wówczas dokonania analizy możliwych sposobów rozwiązania problemu bez ograniczania się do wąskiej klasy narzędzi, następnie wyboru kilku najbardziej obiecujących, za pomocą których podjęta jest próba rozwiązania postawionego problemu. Dla każdego kandydata przeprowadzony jest proces doboru optymalnego zestawu parametrów sterujących algorytmem, analiza wrażliwości algorytmu na niewłaściwy dobór tychże parametrów i ostateczne porównanie z konkurencyjnymi rozwiązaniami. W przypadku algorytmów stochastycznych powyższe analizy powinny mieć charakter statystyczny, przy odpowiedniej liczbie prób dla każdego wariantu rozwiązania.

Autor niniejszej rozprawy do rozwiązania postawionych problemów zastosował znany od początków obecnego wieku algorytm CMA-ES i praktycznie bez żadnych własnych oryginalnych modyfikacji. Dowodzi tezę o możliwości uzyskania określonego sukcesu za pomocą stochastycznego algorytmu optymalizacji przedstawiając jedynie swoiste świadectwo—konkretny znany z literatury przedmiotu algorytm.

Doktorant brał udział w tworzeniu systemów akwizycji danych w eksperymencie NA61 i filtracji danych, oraz oprogramowania SHINE opisanych w rozdziałach 4, 5 i 6 dysertacji. Niestety nie wskazał, jaka była jego rola w licznych zespołach pracujących nad tymi narzędziami i jakie konkretne elementy tych systemów są jego autorstwa. Trudno ocenić, czy wkład doktoranta związany był z profesjonalną umiejętnością implementacji rozwiązań algorytmicznych, czy projektowaniem tych rozwiązań, a jeżeli projektowaniem, to czy na odpowiednim poziomie oryginalności wymaganej od prac doktorskich. Pomijam przy tym fakt braku merytorycznego związku tego wkładu autora pracy do dowodzenia jej tezy.

Przechodząc do analizy dorobku publikacyjnego doktoranta, stosunkowo licznego, bo zawierającego około 20 publikacji, należy zwrócić uwagę, że praktycznie tylko jeden artykuł został opublikowany w czasopiśmie informatycznym—jest to artykuł w czasopiśmie *Schedae Informaticae*, czasopiśmie międzynarodowym, być może niedocenionym punktowo przez MNiSW (11 pkt), wydawanym na macierzystej uczelni doktoranta. Pozostałe artykuły i referaty nie były publikowane w czasopismach i konferencjach informatycznych, nawet te, które przez doktoranta przypisane są do dyscypliny informatyka (a są to *Journal of Physics: Conference Series* i *Journal of Instrumentation*). Znaczna część prac publikowana jest w renomowanych czasopismach, ale nie recenzowanych przez ekspertów dyscypliny *Informatyka*, w szczególności specjalistów w obszarze stochastycznych algorytmów optymalizacji globalnej czy inteligencji obliczeniowej. Dokonania w tych pracach prezentowane bliższe są *Fizyce instrumentalnej*. Według najbardziej elastycznego i liberalnego portalu Google Scholar, indeks Hirsha autora dysertacji wynosi 7 przy ponad 450 cytowaniach. Jednakże niemal wszystkie cytowania i wszystkie artykuły wchodzące do pierwszej siódemki rankingu cytowań, dotyczą artykułów, w których doktorant jest jednym z kilkudziesięciu autorów i trudno ocenić jego wkład w rezultaty badań zawartych w tych publikacjach.

## **Podsumowanie**

Biorąc pod uwagę wymagania stawiane rozprawom doktorskim zawarte w Ustawie o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 r. w art.13 ust 1, po uwzględnieniu całej rozprawy doktorskiej i dorobku publikacyjnego, stwierdzam że:

- a) opiniowana rozprawa doktorska autorstwa Pana mgr Oskara Wyszyńskiego **nie spełnia** wymagań wyżej wspomnianej ustawy;
- b) wnoszę o odrzucenie rozprawy i nie dopuszczenie jej do publicznej obrony w dyscyplinie Informatyka.

W związku z negatywną opinią co do dorobku naukowo-badawczego doktoranta w dyscyplinie Informatyka, w niniejszej opinii nie zamieściłem szczegółowych uwag krytycznych dotyczących zarówno aspektów merytorycznych, jak i edytorskich.

A handwritten signature in dark ink, appearing to be 'A. Janowski', written in a cursive style.