

dr hab. Jarosław Miszczak
Instytut Informatyki Teoretycznej i Stosowanej PAN
Bałtycka 5, 44-100 Gliwice

Gliwice, 25.08.2022 r.

Recenzja pracy doktorskiej
mgra inż. Marka Wróblewskiego
*Kwantowe metody obliczeniowe w hybrydowych klasyczo-kwantowych
systemach rekomendacyjnych*

Przedmiotem recenzji jest rozprawa doktorska *Kwantowe metody obliczeniowe w hybrydowych klasyczo-kwantowych systemach rekomendacyjnych* przygotowana przez mgr. inż. Marka Wróblewskiego pod kierunkiem dra hab. inż. Marka Sawerwaina, prof. UZ.

1 Omówienie pracy

Podjęte przez Autora badania dotyczą nowoczesnych zagadnień z pogranicza teorii informacji, uczenia maszynowego i fizyki. Zaproponowane w pracy rozwiązania wpisują się w nurt współczesnych badań w tej dziedzinie i w mojej opinii pozwalają na identyfikację słabych i mocnych stron hybrydowego podejścia do kwantowego przetwarzania informacji, które jest powszechnie uznawane za główny paradygmat w zakresie projektowania kwantowych systemów obliczeniowych.

Główna część rozprawy składa się z sześciu rozdziałów. Rozdział 1 zawiera streszczenie rezultatów przedstawionych w pracy, w tym główne tezy oraz wykaz opublikowanych prac, będących podstawą rozprawy. Rozdział 2 zawiera krótkie wprowadzenie do teorii obliczeń kwantowych oraz przegląd podstawowych pojęć wykorzystywanych w rozprawie. Dodatkowo, Rozdział 6 stanowi podsumowanie pracy. Rozdziały od 7 do 11 stanowią dodatek do pracy i zawierają przedstawienie podstaw teoretycznych oraz zestawienie szczegółów technicznych związanych z wynikami prezentowanymi w pracy.

Zasadnicze wyniki uzyskane przez Autora są zawarte w Rozdziałach od 3 do 5 i pogrupowane są w dwie części.

Rozdziały 3 i 4 zawierają opis hybrydowego klasyczo-kwantowego systemu rekomendacyjnego oraz mechanizmu profilowania w opisywanym systemie. W Rozdziale 3 opisane są poszczególne etapy budowania systemu, w tym proces kodowania cech i preferencji w stanach kwantowych, organizacja danych w rejestrach kwantowych oraz szczegółowa budowa obwodu kwantowego. Omówione są również elementy takie jak mechanizm pozyskiwania danych oraz całościowa architektura systemu. W Rozdziale 4 wprowadzono i wyjaśniono koncepcję hybrydowego mechanizmu profilowania w systemach rekomendacyjnych, bazującą na hybrydowych drzewach decyzyjnych.

Natomiast Rozdział 5 ma charakter bardziej teoretyczny i przedstawia wyniki związane

z numeryczną analizą splątania w procesie rekomendacyjnym realizowanym w układach opisanych w Rozdziałach 3 i 4. Autor wykorzystuje miarę splątania bazującą na zachowaniu się stanów kwantowych pod działaniem częściowej transpozycji.

2 Wkład autora

Oryginalny wkład Autora polega na zaproponowaniu konstrukcji systemu rekomendacyjnego wykorzystującego kwantowe przetwarzanie danych. Autor przedstawił rozwiązanie kompletne, zawierające wszystkie elementy takiego systemu, począwszy od sposobu pozyskiwania danych, poprzez metody kodowania elementów danych, a kończąc na konstrukcji obwodu kwantowego realizującego procesy rekomendacji i opisie hybrydowej architektury systemu.

Autor zaprezentował również pojęcie hybrydowego mechanizmu profilowania bazujące na koncepcji hybrydowych drzew decyzyjnych oraz wprowadził wykorzystanie podejścia zespołowego poprzez konstrukcje hybrydowych lasów decyzyjnych. Zaproponowane mechanizmy zostały zaimplementowane i przetestowane na komputerach kwantowych reprezentujących obecny stan rozwoju techniki, a dzięki uzyskanym wynikom Autor potwierdził możliwość realizacji hybrydowych kwantowo-klasycznych mechanizmów rekomendacji we współczesnych systemach technicznych.

3 Uwagi krytyczne i pytania

Przedstawiona praca doktorska obejmuje szeroki zakres wiedzy – począwszy od metod kwantowego przetwarzania danych, poprzez systemy pozyskiwania i kodowania danych, a kończąc na metodach uczenia maszynowego. Biorąc pod uwagę tak szeroki zakres tematyczny trudno nie zauważyć, iż pewne elementy są potraktowane pobieżnie i wymagają doprecyzowania. Przedstawiona poniżej lista uwag i pytań zorganizowana jest w odniesieniu do poszczególnych rozdziałów pracy.

Rozdziały 1 i 2

W zakresie systemów metod kwantowego przetwarzania danych brakuje odniesienia do stanu wiedzy oraz omówienia obecnie rozwijanych metod. Autor nie koncentruje się na metodach działających na idealnych komputerach kwantowych w modelu bramkowym. W pracy nie ma odniesienia do algorytmów rozwijanych na potrzeby obliczeń na komputerach z ograniczeniami pamięci i czasu wykonania.

- Jakie są ograniczenia metod przedstawionych w pracy biorąc pod uwagę czynniki takie jak struktura połączeń między kubitami fizycznymi, czas wykonania bramek kwantowych oraz pojemność pamięci kwantowej?
- Czy Autor próbował oszacować jaki jest koszt – np. głębokość obwodu, liczba kubitów – realizacji przedstawionego algorytmu kwantowego dla realistycznej bazy danych? Jak oszacowanie takie ma się do planów rozwoju komputerów kwantowych największych dostawców?

- Proszę o bardziej precyzyjne wyjaśnienie co to znaczy, iż *stosowanie klasycznych metod jest czasowo niewydajne* (str. 20)? Proszę o przedstawienie złożoności obliczeniowej metod, z którymi Autor porównuje swoje rozwiązania.
- Proszę o doprecyzowanie, w jaki sposób dodanie elementu kwantowego (tj. superpozycji) pozwala na ominięcie problemu *zimnego startu* (str. 19)?

Rozdział 3

- W uwadze 3.1, str. 53, Autor stwierdza, iż *proces budowania jest wykonywany tylko raz*. Proszę o doprecyzowanie, czy chodzi tu o projektowanie obwodu czy o wykonanie obwodu na komputerze kwantowym?
- Proszę o doprecyzowanie kilku elementów dyskusji złożoności proponowanego algorytmu.
 - Realizacja części kwantowej (Rys. 3.10) wymaga pomiaru, w wyniku którego uzyskujemy rozkład prawdopodobieństwa. Dlatego do uzyskania wyników konieczne jest wielokrotne wykonanie obwodu. Czy złożoność podana we wzorze (3.2), str. 54, odnosi się do liczby bramek czy do ilości powtórzeń obwodu?
 - Standardową miarą kosztu dla obwodu kwantowego jest skalowanie liczby koniecznych operacji dwukubitowych. Jakie jest skalowanie liczby operacji CNOT dla budowania bazy danych (Rys. 3.12) i kwantowego k-NN (Rys. 3.13)?
- W formule na ilość powtórzeń kwantowego algorytmu k-NN dla $P = 1$ oraz $P = 0$ otrzymujemy $\mathcal{O}(0)$ (str. 58). Proszę o wyjaśnienie tej zależności.
- Implementacja zaproponowanego mechanizmu została udostępniona jako publiczne repozytorium (pozycja [8] w spisie literatury). Takie działanie pozwala na częściową weryfikację przeprowadzonych eksperymentów numerycznych. Niestety w repozytorium brakuje dokumentacji, instrukcji uruchomienia kodu oraz opisu wyników, który umożliwia porównanie odtwarzanych wyników z wynikami Autora.

Rozdział 4

- W przedstawionym mechanizmie profilowania użytkowników Autor bazuje na koncepcji kwantowych drzew decyzyjnych. Nie jest dla mnie jasne, jaki jest oryginalny wkład Autora i jak prezentowane wyniki mają się do innych prac w tym zakresie.
- Pozycje [30] i [85] spisu literatury mają ten sam tytuł i nie jest jasne, do których wyników odwołuje się Autor. Proszę o doprecyzowanie w tym zakresie.
- Kwantowe drzewa decyzyjne były rozważane po raz pierwszy przez Farhiego i Gutmanna (pozycja [51] spisu literatury). W sekcji 4.3.1 Autor powołuje się na tę pracę i stwierdza, iż *las może zostać zbudowany z jednego drzewa bądź z wielu drzew, jednakże opisujących konkretny obiekt*. Proszę o wyjaśnienie, jaki jest związek podejścia zaprezentowanego w pracy z wynikami z pracy [51].
- Klasyczne lasy losowe są bardzo elastyczną metodą, którą można zastosować do szerokiej gamy danych. Jakie są wady i zalety hybrydowych lasów klasyczno-kwantowych w stosunku do lasów klasycznych?

Rozdział 5

- Wyniki w zakresie pomiaru splątania bazują na pakiecie oprogramowania stworzonego przy współudziale Autora. Proszę o sprecyzowanie roli Autora w przygotowaniu tego pakietu.
- Rozdział zawiera szereg wyników związanych z pomiarem splątania, ale nie są one powiązane z obserwacjami dotyczącymi procesu rekomendacji. Czy Autor zaobserwował powiązania między strukturą bazy danych a wartościami ujemności bądź innymi parametrami określającymi nieklasycyzność systemu?

4 Podsumowanie

Przedstawiona praca doktorska podejmuje interesującą, aktualną tematykę i dotyczy zagadnień, które bazują na solidnej znajomości informatyki kwantowej, jak i problematyki przetwarzania danych. Autor uzyskał wyniki, które są pierwszym krokiem w kierunku zastosowania kwantowego przetwarzania danych w hybrydowych systemach uczenia maszynowego.

Silną stroną pracy jest staranne zaprezentowanie próby zastosowania kwantowego przetwarzania danych do realistycznego problemu. Dzięki temu praca wpisuje się w aktualne trendy w dziedzinie informatyki kwantowej, w ramach których duży nacisk kładzie się na budowanie praktycznych rozwiązań opartych na technologiach kwantowych. Należy zaznaczyć, iż zadanie podjęte przez Autora jest ambitne i świadczy o odważnym podejściu do badań w zakresie kwantowego przetwarzania informacji.

W mojej opinii przedstawiona rozprawa jest interesującym krokiem w kierunku praktycznego zastosowania obliczeń kwantowych w technologiach informacyjno-komunikacyjnych. Systemy hybrydowe, bazujące na oddelegowaniu części obliczeń do koprocatora kwantowego, są obecnie głównym paradygmatem rozwoju w tym zakresie. Przedstawiona praca pokazuje, iż systemy tego typu wymagają umiejętnego połączenia dostępnych narzędzi klasycznych z możliwościami komputerów kwantowych. Jednocześnie pozwala ona na wskazanie elementów, które mogą dać przewagę nad systemami klasycznymi. Dzięki temu zaprezentowane wyniki stanowią oryginalny wkład w rozwój informatyki jako dziedziny nauki i techniki.

W związku z powyższym, uważam iż przedstawiona spełnia wymagania określone w *Ustawie o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki*. Na tej podstawie wnioskuję o dopuszczenie Autora rozprawy do dalszych etapów przewodu doktorskiego.



.....
Jarosław Miszczak