

Częstochowa, dn. 04 maja 2022 r.

prof. dr hab. inż. Rafał Scherer
Katedra Inteligentnych Systemów Informatycznych
Wydział Inżynierii Mechanicznej i Informatyki
Politechnika Częstochowska
al. Armii Krajowej 36
42-200 Częstochowa

Recenzja

rozprawy doktorskiej mgr inż. Krzysztofa Fonała, pt.: Algorytmy dekompozycji tensorów w przetwarzaniu danych masywnych.

Promotor: dr hab. inż. Rafał Zdunek, prof. uczelni

Niniejszą recenzję opracowano na wniosek Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Politechniki Wrocławskiej, która mocą uchwały z dnia 16 lutego 2022 roku powołała mnie na recenzenta.

1. Charakterystyka tematu, celu i tezy badawczej rozprawy

Przeważająca większość danych reprezentowana jest w formie macierzowej lub tensorowej. Dane są coraz obszerniejsze i o większej złożoności. Praca dotyczy algorytmów dekompozycji dużych danych wielomodalnych reprezentowanych w formie tensorów. Ogólnym celem pracy było opracowanie algorytmów efektywnej dekompozycji danych wielomodalnych poprzez rozproszenie obliczeń, zmniejszenie złożoności poprzez randomizację oraz umożliwienie aktualizacji przy nadchodzących nowych danych.

2. Zawartość rozprawy

Recenzowana praca mgr inż. Krzysztofa Fonała składa się ze streszczenia, spisu symboli i skrótów, sześciu rozdziałów, dodatku oraz bibliografii. Dokument liczy 166 stron.

Pierwszy rozdział jest wprowadzeniem do reprezentacji macierzowej i tensorowej w informatyce oraz formuluje tezę pracy: *Istnieją strategie obliczeniowe, dzięki którym można zastosować standardowe modele dekompozycji tensorów do przetwarzania danych masywnych.* Podane są również podstawowe definicje: iloczyn Hadamarda Kroneckera, Khatri-Rao, dekompozycja SVD, rozwinięcie względem modu n-tego, iloczyn tensora i macierzy względem modu n, kontrakcja tensora, iloczyn wewnętrzny tensorów, dekompozycje CP, Tuckera czy

WPLYNĘŁO

04-05-2022

RDN-IT/112/2022

ciąg tensorów. Dalej podana jest krótki rys historyczny dekompozycji macierzy. Wprowadzono dekompozycję EVD, SVD, analizę PCA,

Rozdział kończy wykaz dorobku naukowego, składający się z pięciu artykułów w czasopiśmie, dziewięciu artykułów w materiałach konferencji międzynarodowych, oraz rozdziału w monografii.

Rozdział 2 dotyczy rozpraszania obliczeń z ograniczeniami nieujemności, gdzie dane wejściowe i wynikowe mają być nieujemne. W większości zadań dotyczących przetwarzania danych multimedialnych takich jak analiza języka naturalnego, obrazów i wideo mamy do czynienia z danymi nieujemnymi. Rodzina takich obliczeń nazywana jest Nonnegative tensor factorization (NTF), a bazuje najczęściej na dekompozycji CANDECOMP / PARAFAC (CP). Większość metod zaprezentowanych w literaturze zakłada, że nie ma ograniczeń pamięci komputera aby przetwarzać cały zbiór danych, co przy obecnie rozważanych problemach jest często niemożliwe. Autor zaproponował metodę rozproszenia obliczeń opartą na szybkim algorytmie HALS (Hierarchical Alternating Least-Squares). Algorytm ten jest osiąga wyraźnie szybciej zbieżność od konkurencyjnych metod NTF. Na początku rozdziału Autor omawia szczegółowo algorytm HALS. Okazuje się, że algorytm ten jest trudniejszy do modyfikacji w wersji rozproszonej niż inne metody NTF. Doktorant zaproponował dwa podejścia wykorzystujące model rozproszonego programowania MapReduce.

Pierwszym podejściem było podejście algebraiczne, w którym rozpraszany jest jeden element wymagający dużych zasobów. Możliwe jest przez to działanie algorytmu dla danych o dowolnym wymaganiu na pamięć. Zaproponowana modyfikacja, nazwana D-HALS, jest zgodna z algorytmem HALS i skaluje się liniowo. Oczywiście wadą podejścia jest duży narzut czasowy wprowadzany przez model MapReduce.

Drugim podejściem zaproponowanym w rozdziale 2 było podejście geometryczne, w którym cały algorytm HALS jest uruchamiany jest w funkcji Map modelu MapReduce, a wyniki łączone są w funkcji Reduce. Dzięki temu, zaproponowana wersja algorytmu HALS nazwana DG-HALS skaluje się do dowolnych rozmiarów i prędkości, zależnie jedynie od dostępnej mocy obliczeniowej. W podejściu geometrycznym, w odróżnieniu od algebraicznego, występuje mniejsza dokładność działania w porównaniu do oryginalnego algorytmu HALS. Spadek jakości działania jest odwrotnie proporcjonalny do stopnia rzadkości macierzy.

Zaproponowane modyfikacje D-HALS zostały zaimplementowane w Matlab 2016a przetestowane na syntetycznym zbiorze danych oraz popularnym zbiorze MovieLens. Eksperymenty wykonano metodą Monte Carlo, przez uruchomienie 30 prób. Wykazały, że algorytm skaluje się liniowo. Następnie zbadano algorytmy D-HALS-NTF (tu skrót ten pojawia się pierwszy raz) oraz DG-NMF na kilku zbiorach syntetycznych oraz popularnych benchmarkach: MNIST, 20NewsHome, RCV1, COIL-100, CIFAR-10 oraz MovieLens. Eksperymenty wykonano metodą Monte Carlo, przez uruchomienie 10 prób. Badania wykazały, że modyfikacja algebraiczna daje dokładniejsze wyniki, natomiast geometryczna jest szybsza.

Rozdział 3 omawia metody zmniejszania skali problemu, uniezależniające częściowo obliczenia od mocy obliczeniowej. Ponadto, zmniejszona postać danych ma zachować

znaczenie informacji, a oparta jest na idei sieci tensorowych, w postaci Tensor Train oraz Hierarchical Tucker, będącego rozszerzeniem klasycznego modelu Tuckera do wersji drzewiastej. Autor zaproponował metodę dekompozycji tensorów metodą rekurencyjną, nazwaną Fast Hierarchical Tucker Decomposition (FHTD).

Następną modyfikacją było mnożenie tensora przez losową macierz względem dominującego modu, co zmniejsza wymiarowość tego modu, bez znaczącego zmniejszenia ilości informacji, nazwaną Tensor Train Decomposition with Simple Randomness (TTD-SR). Oryginalna macierz jest dekomponowana po zmniejszeniu jej przez projekcję z macierzą losową. Eksperymenty wykazały możliwość zmniejszania rozmiaru danych o dwa rzędy wielkości bez znaczącej utraty jakości. Ostateczna forma algorytmu została nazwana Fast Hierarchical Tucker Decomposition with Simple Mode Preservation and Randomization (FHTDSMPR).

Eksperymenty w rozdziale 3 wykonano na zbiorach The University of Iowa Musical Instrument Samples (MIS) zawierającym próbki nagrań instrumentów muzycznych, ORL ze zbiorami twarzy czterdziestu osób, SPhone z sygnałami czujników ruchu ze smartfonów oraz Columbia Object Image Library (COIL-100) z obrazami 100 obiektów na czarnym tle. Zbiory zostały poddane augmentacji wraz z rozszerzeniem do pięciomodalności. Zaproponowane algorytmy FHTDSMPR oraz TTD-SR zostały porównane z wybranymi metodami występującymi w literaturze jak PCA, HALS-NMF, HOSVD, HOCMM, HALS-NTF czy TT-SVD. Eksperymenty były uruchamiane 50 razy (Monet Carlo + walidacja krzyżowa). Eksperymenty wykazały bardzo dobrą jakość klasyfikacji, przy jednoczesnym znaczącym zmniejszeniu rozmiar danych. Nie mogłem doszukać się jaki algorytm klasyfikacji zastosował Autor.

Rozdział 4 dotyczy przypadków, w których nie dysponujemy całością zbioru danych, a dane przyrastają w sposób ciągły i należy modyfikować na bieżąco stan systemu. Przetwarzanie wsadowo powodowałoby konieczność przeliczania całego zbioru danych po otrzymaniu nowej porcji danych. Autor zaproponował dwa podejścia do przetwarzania takich danych strumieniowych, oba oparte o strumieniową dekompozycję Tuckera z ograniczeniem na nieujemność (NTD). W zaproponowanym modelu o nazwie Incremental NTD dane są sekwencjami podtensorów w domenie czasu lub numeru próbki. Drugie podejście jest wzorowane na metodzie Kaczmarz-online, pozwalające estymację czynników dynamicznych, których charakterystyka zmienia się w czasie.

Testy wykonano na trzech syntetycznych zbiorach danych. Zaproponowane metody BK-NTD i RI-NTD zaimplementowano w środowisku MATLAB 2016b i porównano z algorytmem HALS-NTD istniejącym w tym środowisku. Autor zastosował analizę Monte Carlo, inicjalizując eksperymenty 10 razy. Algorytmy są szybsze o rząd i dwa rzędy wielkości od algorytmu HALS-NTD, mogąc przy tym przetwarzać dane o dowolnej długości.

Rozdział 5 prezentuje sposób łączenia danych z wielu tensorów oraz metodę uzupełniania danych. Algorytm uzupełniania danych został przetestowany na syntetycznym zbiorze danych powstałym przez usuwanie pikseli z kolorowego obrazu o rozmiarach 512×512 . Algorytm rekonstruował obraz lepiej niż konkurencyjne algorytmy z tej rodziny algorytmów.

Rozdział 6 jest podsumowaniem rozprawy, w którym zebrano w jednym miejscu wnioski kończące opisy poszczególnych autorskich metod i eksperymentów. Autor podał również

przyszłe kierunki badawcze jakie ma zamiar podjąć: badania rzeczywistych wielkich zbiorów danych, porównanie i osadzenie proponowanych metod w świecie uczenia maszynowego, szczególnie sieci neuronowych, oraz zastosowanie dekompozycji tensorowych do kompresji sieci neuronowych.

Dalej następuje dodatek z kodami źródłowymi.

Pracę kończy bibliografia składająca się z 173 aktualnych pozycji.

Ogólnie, zasadnicze i oryginalne rezultaty pracy można podsumować następująco:

- Opracowanie wprowadzenia do tematyki i przeglądu literatury dotyczącej nieujemnej dekompozycji dużych tensorów.
- Stworzenie algebraicznej i geometrycznej metody rozpraszania obliczeń nieujemnej faktoryzacji tensorów.
- Opracowanie metod zmniejszania skali problemu oparte na sieciach tensorowych i hierarchicznym modelu Tuckera.
- Stworzenie algorytmów estymacji czynników dla danych zmieniających się w czasie, tzw. strumieniowych,
- Opracowanie metod łączenia cech z wielu tensorów i uzupełniania danych w tensorach,
- Przeprowadzenie bardzo dobrze zaplanowanych eksperymentów.

Mgr Fonał opublikował ponadprzeciętną liczbę prac naukowych: składający się z pięciu artykułów w czasopiśmie, dziewięciu artykułów w materiałach konferencji międzynarodowych, oraz rozdziału w monografii. Zaprezentowany materiał pokazuje, że Doktorant zrealizował cel pracy.

3. Uwagi krytyczne i wskazówki dotyczące rozprawy

Praca napisana jest przejrzysto z czytelnymi rysunkami oraz algorytmami. Eksperymenty wykonane na dużych zbiorach danych (w miarę możliwości) wykazały bardzo dobre działanie zaproponowanych metod. Poniżej zamieszczam kilka pytań i wątpliwości, które zrodziły się w czasie czytania pracy:

- Praca porównuje autorskie rozwiązania z podobnymi metodami. Ciekawym byłoby porównanie tych metod pod względem szybkości, zapotrzebowania na pamięć oraz jakości działania z metodami nieliniowymi, np. głębokimi sieciami neuronowymi, np. sieciami w pełni splotowymi do rekonstrukcji obrazu.
- Nie mogłem znaleźć sposobu klasyfikacji w badaniach w podrozdziale 3.4. Można domyślać się użycia algorytmu kNN, wzmiankowanego wcześniej, ale nie jest to jasne. Nie ma również podanych parametrów algorytmu.
- str. 91: Typy algorytmów na rysunku 3.4 nie są opisane w sposób zrozumiały dla czytelnika.
- Skrót D-HALS-NTF pojawia się pierwszy raz na stronie 49 i nie jest wyjaśniony.
- Skrót IC-TT nie jest chyba wyjaśniony, a użyty w algorytmie 24.

- Tytuł rozdziału 4 odnosi się do terminu czas rzeczywisty, lecz może ściślejszym terminem byłoby użycie terminu odnoszącego się do danych strumieniowych, przeciwieństwa przetwarzania wsadowego.

4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że Pan mgr inż. Krzysztof Fonał w rozprawie doktorskiej „Algorytmy dekompozycji tensorów w przetwarzaniu danych masowych”:

- Zrealizował cel rozprawy,
- Opracował wprowadzenie do tematyki i przegląd literatury dotyczącej nieujemnej dekompozycji dużych tensorów,
- Stworzył algebraiczną i geometryczną metodę rozpraszania obliczeń nieujemnej faktoryzacji tensorów.
- Stworzył metody zmniejszania skali problemu oparte na sieciach tensorowych i hierarchicznym modelu Tuckera.
- Stworzył algorytmy estymacji czynników dla danych zmieniających się w czasie, tzw. strumieniowych,
- Opracował metody łączenia cech z wielu tensorów i uzupełniania danych w tensorach,
- Przeprowadził serię bardzo dobrze zaplanowanych eksperymentów,
- Wykazał się umiejętnością samodzielnej pracy badawczej, znajomością literatury światowej i wiedzą w zakresie wielomodowych obliczeń wielkich danych,
- Zadbał o popularyzację wyników swoich badań w wysokopunktowanych wydawnictwach.

Recenzowana praca spełnia wymagania ustawy o tytule i stopniach naukowych w dyscyplinie naukowej Informatyka Techniczna i Telekomunikacja. Wnoszę o jej przyjęcie i dopuszczenie do dalszych etapów postępowania doktorskiego. Jednocześnie ze względu na wysoki poziom naukowy rozprawy, wnioskuję o wyróżnienie rozprawy.

Robert Scherer