

---

Dr hab. inż. Wojciech Kotłowski, prof. PP  
Instytut Informatyki  
Politechnika Poznańska  
ul. Piotrowo 2, 60-965 Poznań  
tel: (+48) 61 665 2936  
wkotlowski@cs.put.poznan.pl

---



Poznań, dn. 10 września 2023 r.

## Recenzja rozprawy doktorskiej

mgr. inż. Piotra Bielaka

*Methods for selected problems in unsupervised graph representation learning*

*(Metody dla wybranych problemów nienadzorowanego uczenia reprezentacji grafów)*

### 1 Problem badawczy jego znaczenie

Rozprawa doktorska mgr. inż. Piotra Bielaka dotyczy problemów uczenia maszynowego dla danych o strukturze grafu. Takie dane występują coraz częściej we współczesnych zastosowaniach – można, za Autorem pracy, podać przykłady badania sieci społecznościowych, chemii obliczeniowej i relacji między atomami w cząsteczkach, czy analizy sieci cytowań artykułów naukowych. Grafy opisują obiekty (węzły) wraz z relacjami zachodzącymi między nimi (krawędzie). W ogólności, węzły i krawędzie zawierają dodatkowe informacje w postaci przypisanych im wektorów cech. Głównym celem pracy jest uczenie się reprezentacji grafów, czyli utworzenia wektorów osadzenia/zanurzenia dla wszystkich węzłów bądź krawędzi grafu w przestrzeni o mniejszym wymiarze, niż oryginalna przestrzeń danych wejściowych. Reprezentacja taka powinna brać pod uwagę wzajemne relacje między elementami grafu, które są kluczowe dla uzyskania zadowalającej trafności predykcji w późniejszych zadaniach uczenia maszynowego, takich jak np. klasyfikacja węzłów bądź całego grafu, czy przewidywania powiązań między węzłami. Z powodu rosnącego zainteresowania problemami grafowymi w przemyśle i nauce, temat pracy Doktoranta wydaje się świetnie trafiać we współczesne trendy badawcze i może prowadzić do istotnych praktycznych zastosowań.

Tematyka uczenia się reprezentacji grafów nie jest oczywiście nowa, a wiele podejść, szczególnie tych opartych na głębokich sieciach neuronowych, rozwijanych jest w ostatnich

WPLYNĘŁO

12-10-2023

ROTIiT/190/2023

*Kotłowski*

latach bardzo intensywnie. Warto tu na pewno wspomnieć o najbardziej znaczącym kierunku badań, czyli grafowych sieciach neuronowych (GNN), które w swojej podstawowej wersji opierają się na przekazywaniu w danej warstwie informacji od bezpośrednich sąsiadów w grafie, ich agregacji i aktualizacji na tej podstawie wewnętrznego stanu węzłów. Poprzez zastosowanie kilku warstw możliwe jest w ten sposób przekazywanie informacji od obszernego sąsiedztwa w grafie, co pozwala późniejszej reprezentacji lepiej uchwycić relacje między węzłami. Pomimo wielu dobrze przeanalizowanych metod i zagadnień uczenia się grafów, Doktorantowi udało się znaleźć szereg problemów, które nie zostały dotąd przebadane w wystarczającym stopniu. W szczególności, skupił się on na metodach *uczenia bez nadzoru*, w których na etapie konstrukcji reprezentacji nie mamy dostępu do informacji na wyjściu (np. etykiet). Jest to z jednej strony ogromna zaleta metod nienadzorowanych, ponieważ nie ma potrzeby etykietowania danych, a jest to zadanie często bardzo kosztowne i czasochłonne, wykonywane przez człowieka. Z drugiej strony, proces trenowania algorytmów nienadzorowanych jest bardziej skomplikowany, ponieważ nie mają one dostępu do informacji zwrotnej na wyjściu, stąd muszą opierać się na zastępczych funkcjach celu (funkcjach błędu), umożliwiającym utworzenie wartościowych reprezentacji. Konstrukcja takich metod jest więc zadaniem trudnym i z pewnością wystarczająco ambitnym, aby stało się tematem rozprawy doktorskiej.

Mgr inż. Piotr Bielak postawił w rozprawie następujące cele badawcze:

1. Weryfikacja wykorzystania korelacji krzyżowej jako miary jakości reprezentacji do uczenia samo-nadzorowanego grafowej sieci neuronowej ze zmodyfikowanych (przez losowe perturbacje) widoków grafu atrybutowanego, tak aby sieć dostarczyła reprezentację lepszą od metod istniejących.
2. Opracowanie modelu grafowej sieci neuronowej celem wyznaczenia dobrej jakościowo (względem metod istniejących) reprezentacji krawędzi w grafie.
3. Konstrukcja przyrostowej metody uczenia się reprezentacji w grafach dynamicznych o niższej (od rozwiązań konkurencyjnych) złożoności czasowej i pamięciowej, a równocześnie o wyższej docelowej jakości.
4. Opracowanie metody umożliwiającej połączenie istniejącej już strukturalnej reprezentacji węzłów z informacją o ich atrybutach, celem uzyskania nowej reprezentacji o jakości wyższej, niż każdy ze składników z osobna.

W mojej opinii, powyższe cele są bardzo dobrze uzasadnione, interesujące, o dużym wymiarze praktycznym, a przede wszystkim nowatorskie i ambitne. Każdy z tych celów doprowadził do oryginalnych i istotnych wyników badawczych opublikowanych w czasopiśmie dziedziny.

## 2 Ocena struktury i zawartości pracy

Recenzowana rozprawa jest napisana czytelnie bardzo dobrym językiem angielskim. Liczy 118 stron i składa się z czterech rozdziałów, poprzedzonych streszczeniem w języku angielskim i polskim. Ma charakter syntezy treści pięciu powiązanych ze sobą tematycznie artykułów naukowych, zbiorczo załączonych jako treść czwartego rozdziału, natomiast trzy poprzedzające rozdziały to wprowadzenie do tematyki pracy, określenie hipotez badawczych, opisanie wyników, oraz podsumowanie. Całościowo strukturę pracy oceniam całkowicie pozytywnie, a podział na rozdziały wydaje się zupełnie naturalny dla pracy opartej na publikacjach.

Rozdział pierwszy ma charakter wprowadzenia. Nakreślona w nim została motywacja podjęcia badań w tematyce pracy, związana z rosnącym zainteresowaniem danymi o strukturze grafów. Opisane zostały grafy atrybutowane, i związane z nimi podstawowe zadania uczenia maszynowego, takie jak klasyfikacja węzłów, krawędzi, przewidywanie połączeń, rekonstrukcja grafów, czy grupowanie wierzchołków. Doktorant opisał również podejścia do uczenia się z danych grafowych, szczegółowo opisując uczenie się reprezentacji. Dużo miejsca zostało poświęcone współczesnym metodom uczenia maszynowego opartych na grafowych sieciach neuronowych. W dalszej części rozdziału wymienione zostały obszary badawcze, które nie zostały w wystarczającym stopniu przeanalizowane w literaturze. Dla każdego ze zidentyfikowanych obszarów Autor postawił cele badawczy z wyżej wymienionej listy. Rozdział jest bardzo czytelny, a postawione cele dobrze umotywowane.

W rozdziale drugim wymienione zostały hipotezy badawcze wraz z wynikami prac, wpięrow w formie krótkiej listy podsumowującej każdą z hipotez (wraz z listą publikacji składających się na pracę doktorską), a następnie omówione po kolei dużo dokładniej. W każdym przypadku wskazana została publikacja bądź publikacje związana z danym celem i opisane zostały szczegółowo wyniki w niej zawarte. Uważam układ tego rozdziału za wręcz wzorcowy. Pozwala on czytelnikowi zapoznać się z treścią pracy, przedstawiając główne założenia i wyniki, co znacznie ułatwia późniejsze czytanie samych publikacji. Każda z publikacji jest jednoznacznie przyporządkowana do danego zadania badawczego, co umożliwia łatwą weryfikację wypełnienia wszystkich celów założonych w pracy, a równocześnie potwierdza spójność tematyczną całego zbioru prac.

Rozdział trzeci zawiera wnioski z prac badawczych i podsumowanie wyników.

## 3 Ocena wkładu oryginalnego

Rozprawa jest oparta na kilku artykułach, których współautorem jest Doktorant:

1. Piotr Bielak, Tomasz Kajdanowicz, Nitesh V. Chawla: *Graph Barlow Twins: A self-supervised representation learning framework for graphs*. Knowledge-Based Systems, vol. 256, 2022

2. Piotr Bielak, Tomasz Kajdanowicz, Nitesh V. Chawla: *AttrE2vec: Unsupervised attributed edge representation learning*. Information Sciences, vol. 592, 2022
3. Piotr Bielak, Kamil Tagowski, Maciej Falkiewicz, Tomasz Kajdanowicz, Nitesh V. Chawla: *FILDNE: A Framework for Incremental Learning of Dynamic Networks Embeddings*. Knowledge-Based Systems, vol. 236, 2022
4. Piotr Bielak, Daria Puchalska, Tomasz Kajdanowicz: *Retrofitting Structural Graph Embeddings with Node Attribute Information*. Computational Science – ICCS 2022, vol. 13350, Lecture Notes in Computer Science, Springer, 2022
5. Piotr Bielak, Jakub Binkowski, Albert Sawczyn, Katsiaryna Viarenich, Daria Puchalska, Tomasz Kajdanowicz: *A deeper look at Graph Embedding RetroFitting*. Journal of Computational Science, vol. 68, 2023

Pierwsze trzy prace opublikowane są w czasopismach wycenianych na 200 punktów MEIN każde, o współczynniku *Impact Factor* (IF) powyżej 8. Kolejne dwie prace mają, odpowiednio 140 i 100 punktów MEIN. Sumaryczna liczba punktów MEIN w powyższym dorobku jest więc równa 840, a sumaryczny IF przekracza wartość 28, co jest osiągnięciem imponującym jak na zaledwie dwuletni rozkład prac. Dorobek uznaję za więc za bardzo wartościowy, z nawiązką spełniający wymogi do finalizacji przewodu doktorskiego.

Rozprawa zawiera wiele oryginalnych i nowatorskich wyników, które zostały już pozytywnie zweryfikowane i docenione przez recenzentów na etapie przyjmowania prac do czasopism. Poniżej przedstawiam ich podsumowanie:

- Propozycja modelu *Graph Barlow Twins* do samo-nadzorowanego uczenia się reprezentacji. Pomysł zbudowania funkcji celu na macierzy korelacji krzyżowej perturbowanych losowo grafów jest całkowicie nowatorski w kontekście uczenia się reprezentacji w danych o charakterze grafowym, a zarazem świetnie tu pasuje z uwagi na wspomniane przez Autora trudności ze zdefiniowaniem próbkowania negatywnego. Uzyskane w ten sposób algorytmy wykazują się bardzo wysoką trafnością predykcyjną w różnorodnych zadaniach docelowych, pokonując istniejące rozwiązania.
- Wprowadzenie metody wyznaczania reprezentacji dla krawędzi poprzez zastosowanie metody uczenia kontrastowego i agregację cech pochodzących od wierzchołków tworzących krawędź i ich sąsiadów. Doktorant trafnie użył tu spaceru losowego, pozwalającego na bezbolesne skalowanie metody do dużych grafów. Zaproponowany został szereg metod agregacji, w tym również metody „nauczalne”, takie jak rekurencyjna sieć GRU. Uzyskany w ten sposób algorytm, *AttrE2vec*, okazuje się budować reprezentacje krawędzi o wysokiej jakości, potwierdzone w bardzo obszernym eksperymencie obliczeniowym na wielu zbiorach danych i zadaniach docelowych. Doktorant poświęcił dużo pracy, aby włączyć do eksperymentu wiele algorytmów porównawczych (*baselines*), tym silniejsze stają się więc wnioski z badań wskazujące na przewagę metody *AttrE2vec*.

- Konstrukcja metod uczenia przyrostowego *FILDNE* dla grafów dynamicznych, tzn. grafów zmieniających się w czasie. Zaproponowany algorytm tworzy finalną reprezentację na podstawie reprezentacji uzyskiwanych z kolejnych „migawek” grafu (stanów w poszczególnych chwilach czasowych), wraz z ich dopasowaniem przy użyciu narzędzi algebry liniowej, i finalnej kombinacji wypukłej, ze współczynnikami, które można dopasować używając metod strojenia hiperparametrów (np. przeszukiwanie na siatce, lub metody bayesowskie). Również tym wynikom towarzyszy obszerny eksperyment obliczeniowy na kilku zadaniach docelowych i siedmiu rzeczywistych zbiorach danych. Zaproponowana metoda nie tylko redukuje zasoby obliczeniowe i pamięciowe, ale również prowadzi do reprezentacji o bardzo wysokiej jakości na zadaniach docelowych.
- Propozycja metody połączenia ze sobą informacji o atrybutach węzłów z istniejącą już strukturalną reprezentacją węzłów, która prowadzi do algorytmu *GERF* i jego ulepszonej wersji *GERF++*. Algorytmy tworzą nową reprezentację, zachowującą informację zakodowaną w reprezentacji strukturalnej, a zarazem wzbogacającą ją o informacje pochodzące z atrybutów. Finalna reprezentacja okazała się w eksperymencie obliczeniowym lepsza od oryginalnej reprezentacji strukturalnej, uczenia się na bazie atrybutów grafu, czy naiwnych metod łączenia obu źródeł informacji.

Każdy z wymienionych wyników prowadzi więc do propozycji nowatorskiej metody rozwiązującej obrany problem badawczy, każdorazowo dokładnie przetestowanej w obszernych eksperymentach obliczeniowych, z których wyłania się jako lepsza bądź co najmniej konkurencyjna względem rozwiązań bazowych.

Warto też na koniec podkreślić, że istotną zaletą większości z zaproponowanych metod jest ich elastyczność względem użycia bazowego głębokiego modelu tworzenia reprezentacji (tzn. grafowego *encodera*) – mogą one pracować z sieciami o potencjalnie dowolnej architekturze, ponieważ *encoder* jest odrębnym, niezależnym modułem w obrębie proponowanych rozwiązań. Ta elastyczność pozwala łatwo dostosować utworzone metody do specyfiki problemu i danych.

Krótko podsumowując, uznaję, że wymienione na wstępie pracy cele udało się Doktorantowi w pełni osiągnąć.

## 4 Dalsze uwagi

Nie kwestionując wartości całościowych wyników zawartych w rozprawie, chciałbym zgłosić poniżej kilka uwag, głównie w formie pytań bądź dyskusji.

- Wykorzystanie macierzy korelacji krzyżowej w metodzie *Graph Barlow Twins*: macierz korelacji jest normalizowana odchyleniami standardowymi i wydaje się trochę mniej stabilnym celem nauki, niż np. zwykła macierz kowariancji krzyżowej, tzn. bez normalizacji. Wydaje mi się, że same ograniczenia trzymania diagonalnych współczyn-

ników w okolicy 1, a poza-diagonalnych w okolicy 0 powinno stanowić wystarczający nacisk na algorytm do stabilizacji wariancji poszczególnych cech, a równocześnie utrzymać współrzędne wektora reprezentacji na jednolitej skali. Czy Doktorant próbował używać tego podejścia, optymalizując po prostu kowariancję? Można by nawet pójść krok dalej i nie odejmować średnich, wtedy funkcję celu można by zapisać w postaci iloczynów skalarnych cech liczonych po batchu.

- W algorytmie AttrE2vec negatywne przykłady do błędu kontrastowego próbkowane są spośród wszystkich węzłów nie odwiedzonych przez spacer losowe. Ale przez losowość samego spaceru zdarza się, że węzły bliskie danej krawędzi trafią do przykładów negatywnych, a dalsze (ale odwiedzone przez spacer) – do pozytywnych. Czy zjawisko takie nie utrudnia uczenia się sieci?
- W metodach *FILDNE* i *k-FILDNE* nie było dla mnie jasne, czy współczynniki mieszania  $\alpha$  (bądź  $\alpha$ ) zależą od chwili czasowej  $t$  (czyli są przy każdej kolejnej „migawce” grafu liczone od nowa), czy są ustalone dla całego „strumienia” grafów. Po zapoznaniu się z publikacją wydaje się, że liczone są każdorazowo przy nowym grafie  $G_t$ . Czy bayesowska metoda estymacji, wymagająca każdorazowo wytrenowania  $k$  regresji logistycznych, nie jest zbyt czasochłonna i nie dominuje pod względem czasu obliczeń pozostałych kroków algorytmu?
- Z równań (2.5) i (2.6) wnioskuję, że w metodach *GERF* i *GERF++*, funkcja celu jest kwadratowa względem nowych wektorów reprezentacji. Oznacza to, że można znaleźć minimum tej funkcji (czyli nową macierz reprezentacji  $\mathbf{Z}^*$ ) w sposób *analytyczny*:  $\mathbf{Z}^* = \mathbf{AZ}$  dla pewnej macierzy  $\mathbf{A}$  o rozmiarach  $n \times n$ , uzyskanej z przyrównaniu gradientu  $\nabla \mathcal{L}(\mathbf{Z}^*)$  do zera.
- Ciekawi mnie, czy Doktorant rozważał publikację wyników badań na głównych konferencjach uczenia maszynowego, takich jak *Neural Information Processing Systems (NeurIPS)*, *International Conference on Learning Representations (ICLR)* czy *International Conference on Machine Learning (ICML)*. Mam wrażenie, że jakość zaprezentowanych prac jest na tyle wysoka, a tematyka na tyle interesująca, że z dużą szansą mogłyby one zostać na takie konferencje przyjęte. Nie ujmuję tu nic bardzo silnym czasopismom, w których Doktorant opublikował pracę, ale wspomniane konferencje mają bardzo dużą „widoczność” w społeczności uczenia maszynowego.

## 5 Konkluzja końcowa

Rozprawę oceniam bardzo dobrze, co w powyższej recenzji podkreśliłem wielokrotnie. Problemy badawcze, z którymi zmierzył się mgr inż. Piotr Bielak, są ambitne i istotne dla postępu w dziedzinie uczenia się reprezentacji danych o charakterze grafów. Sama rozprawa charakteryzuje się bardzo wysokim poziomem merytorycznym i zawiera wiele interesujące rezultatów, jednoznacznie potwierdzających użyteczność i praktyczność zaproponowanych

algorytmów. W mojej opinii Doktorant wykazał się znakomitymi umiejętnościami prowadzenia badań naukowych. Wszystkie postawione w rozprawie cele zostały osiągnięte.

W związku z tym rozprawę oceniam jako spełniającą wymogi stawiane pracom doktorskim i wnoszę o dopuszczenie mgr. inż. Piotra Bielaka do dalszych etapów przewodu doktorskiego, równocześnie proponując wyróżnienie pracy.

Wojciech Kotłowski

dr hab. inż. Wojciech Kotłowski