



Prof. dr hab. Włodzisław Duch  
Katedra Informatyki Stosowanej  
i Laboratorium Neurokognitywne  
Uniwersytet Mikołaja Kopernika, Toruń

Wpłynęło dnia 2022-11-03



Recenzja rozprawy doktorskiej mgr Joanny Dudy-Goławskiej,  
„Analiza ERP z wykorzystaniem wielokanałowego algorytmu Matching Pursuit”.

Mgr Duda-Goławska przygotowała rozprawę doktorską w Zakładzie Fizyki Biomedycznej Instytutu Fizyki Doświadczalnej Wydziału Fizyki Uniwersytetu Warszawskiego, pod opieką dr hab. Jarosława Żygierewicza, profesora UW. Analiza zapisów elektroencefalograficznych, niestacjonarnych szeregów czasowych, jest zagadnieniem bardzo trudnym. Jedną z najczęściej stosowanych metod jest analiza potencjałów wywołanych (Evoked Response Potentials, ERP). Bodźce prezentowane wielokrotnie wywołują reakcje sieci neuronowych mózgu, które po uśrednieniu są w miarę stabilne, pokazując charakterystyczne składowe, podlegające interpretacji. Istnieje wiele metod analiz takich potencjałów, znajdujące liczne praktyczne zastosowania pakiety oprogramowania, takie jak np. ERPLab. Trudno tu zaproponować nowe, oryginalne rozwiązania, a jednak właśnie temu zagadnieniu poświęcona jest praca doktorska mgr Dudy-Goławskiej. Rozprawa, napisana w języku polskim, jest nieznacznie rozszerzoną wersją publikacji Duda-Goławska, Imbir i Żygierewicz, ERP Analysis Using a Multi-Channel Matching Pursuit Algorithm, opublikowanej w czasopiśmie Neuroinformatics, 1-36, 2022.

Metody analiz ERP opierają się na obserwacjach fenomenologicznych, porównaniach komponentów sygnału powstającego w wyniku odpowiedzi na powtarzający się bodziec po określonym czasie z warunkami eksperymentów. Komponenty powinny być powiązane z jakimiś procesami mentalnymi, ale jest to złożona relacja. W pracy doktorskiej jest tylko bardzo skrócony opis zagadnień związanych z analizą i interpretacją ERP. Wiele zagadnień, takich jak analiza informacji z pojedynczych prób (single-trial ERP), nie zostało w niej poruszonych. W kilku miejscach autorka odwołuje się do książki „Applied Event-Related Potential Data Analysis” Stevena Lucka (darmowa zaktualizowana wersja tej książki jest dostępna w LibreText), w której można znaleźć szczegółowe informacje.

Sygnały ERP reprezentują sumaryczną aktywność wielu sieci realizujących różne procesy zachodzące w danym czasie. Klasyczna interpretacja komponentów, jak napisał Luck (2005),

jest „mglista” (nebulous). Kramer (1985) pokazał jednak, że eksperci potrafią prawidłowo rozróżniać ukryte komponenty w przebiegach ERP. Powinno się więc dać opracować metody pozwalające na głębszą interpretację. Kształt przebiegów czasowych ERP jest wynikiem sumowania wkładów od różnych ukrytych komponentów, wywołujących szybko zmieniające się rozkłady czasoprzestrzenne potencjałów na powierzchni głowy. Nie mamy modeli parametrycznych, pokazujących wpływ aktywizacji różnych obszarów mózgu na mierzone potencjały pod wpływem różnych bodźców. Ukryte komponenty powiązane z aktywacjami specyficznej aktywności określonych obszarów mózgu próbujemy odkryć analizując sygnały EEG. Praca mgr Dudy-Goławskiej poświęcona jest zastosowaniu wielokanałowego algorytmu matching pursuit (ang. Multivariate Matching Pursuit, MMP), metod rozwijanych przez neuroinformatyków z Uniwersytetu Warszawskiego, do odkrywania ukrytych komponentów na podstawie analizy sygnałów ERP. Literatura ten temat takich metod w zastosowaniu do potencjałów wywołanych jest bardzo skąpa, znalazłem tylko kilka mało cytowanych publikacji konferencyjnych Selin Aviyente (2006, 2010, 2020, 2022).

W rozdziale drugim Autorka przedstawiła podstawowe metody, z których korzystała w dalszej części pracy. Jest tu model addytywny, wiążący ukryte komponenty z obserwowanym sygnałem, przegląd metod eksploracyjnych i opis algorytmu dopasowania kroczącego. Metody eksploracyjne potraktowane zostały bardzo pobieżnie, wspomniane zostały mikrostany i globalna moc pola (GFP), analiza składowych głównych PCA i składowych niezależnych ICA. Nie ma jednak wzmianki o wariantach ICA, a analiza przestrzenno-czasowa ICA jest bliska celem tego doktoratu i może dostarczyć naturalnych funkcji do bazy atomów. Metody dekompozycji używane są również w analizie falkowej, rozkładzie sygnału na mody empiryczne (EMD), aproksymantach Pade i oscillonach, a więc metodach, których wady i zalety warto było porównać z MMP.

Warto też dokładniej przedyskutować sensowność poczynionych założeń, przedstawionych w rozdziale trzecim, opisującym szczegóły zastosowanej metody. Propagacja zmiany potencjału pola elektrycznego rozchodzi się z prędkością światła, ale zanim neurony się pobudzą i zsynchronizują z innymi grupami neuronów, dając silny sygnał odróżnialny od szumu, mija sporo czasu, co widać na zmianach amplitudy potencjałów wywołanych. Chcemy badać morfologię sygnału pochodzącego od jednocześnie aktywnych populacji, a to nie dzieje się w sposób natychmiastowy. Czy pominięcie opóźnień czasowych jest na pewno uzasadnione?

Zakłada się też, że taki sygnał będzie rozpoznawalny we wszystkich kanałach EEG, ale skoro w pamięci roboczej możemy mieć kilka niezależnych stanów, to będzie tak jedynie w warunkach eksperymentalnych, w których pojedynczy stan dominuje. Na aktywność spoczynkową wpływa wiele konkurujących ze sobą procesów. W pracy doktorskiej analizowane jest konkretnie zadanie, więc w krótkich odcinkach czasu, w których analizujemy ERP, może dominować jeden proces. Jednak powszechnie stosowane założenie, że każdy atom można opisać za pomocą jednego lub dwóch źródeł aktywności w postaci równoważnych dipoli, nie zgadza się z naszą wiedzą o rozległych sieciach. W neuroobrazowaniu przy użyciu fMRI widać zwykle kilka aktywnych obszarów w różnych częściach mózgu.

Rozdział czwarty, zdecydowanie najdłuższy, opisuje testy metody dla symulowanych danych, jak i na danych z badania kategoryzacji emocjonalnej słów. Konstrukcja danych symulowanych opiera się na kombinacji 4 komponentów N100, P2, P300 i LPC (Autorka raz pisze P2, P3, a innym razem P200, P300), modelowanych za pomocą funkcji Gabora. Współczynniki określające względne amplitudy dla poszczególnych komponentów zostały dobrane dla trzech warunków eksperymentalnych, a same topografie dla 19 kanałów EEG widoczne są na kilku rysunkach. Symulacje dla 40 powtórzeń każdego eksperymentu i 50 osób różnią się latencją komponentów i zmiennością amplitudy, zarówno dla różnych osób jak i powtórzeń dla tej samej osoby. Dodano też jako tło niezwiązane z bodźcem fragmenty spoczynkowego EEG. Dane uśredniono po próbach dla badanej osoby oraz po osobach dla danego warunku eksperymentalnego, otrzymując przykładowe sygnały ERP na każdej z elektrod (Rys. 4.4). Metoda MMP prawidłowo wykryła obecność 4 ukrytych atomów składających się na obserwowane komponenty, a ich porównanie z wagami użytymi w symulacjach (Tab 4.2) pokazało dużą zgodność. Potwierdza to możliwości wykrycia ukrytych komponentów metodą MMP.

Dane eksperymentalne pochodzą z badania kategoryzacji emocjonalnej słów, w którym mgr Duda-Goławska brała udział. W tym eksperymencie, opublikowanym w 2019 roku, badani mieli oceniać czy prezentowane słowo jest neutralne, czy można je uznać za emocjonalne. Z bazy 4920 słów opublikowanej przez Kamila Imbira wybrano 135 słów. Każde z nich ma przypisany Znak i Genezę. Dla uproszczenia zredukowano wartości tych parametrów z 9 do 3 poziomów intensywności. Z Rys. 4.2 wynika, że nie były to słowa o skrajnych różnicach w zakresie Znaku (czyli pozytywnych lub negatywnych ocen), za to uwzględniono słowa o ekstremalnych różnicach Genezy. Na tym rysunku skala  $[-2, 1.5]$  i  $[-5, +4]$  nie wygląda na 3 stopnie, chociaż wybrane słowa pewnie się w trzystopniowej skali mieszczą. Zależność amplitudy od 9

kombinacji wartości (Znak, Geneza) oceniano za pomocą średnich z pomiarów EEG, średnich dla komponentów z rekonstrukcji metodą MMP i pojedynczych atomów, które otrzymywane są z analizy skupień dla wszystkich badanych.

Podrozdział 4.4 opisuje szczegółowe wyniki analizy danych eksperymentalnych. Wyniki behawioralne zostały przedstawione bez próby interpretacji ich sensu. Taka jednoznaczna interpretacja nie jest obecnie możliwa, ale warto zadać sobie kilka pytań. Cemu słowa negatywne związane są z dłuższymi czasami reakcji, zwłaszcza dla genezy automatycznej? Wygląda na to, że przypisano automatyczną genezę do słów, których ocena wymaga dłuższego namysłu, takich jak „zakochanie”, wymagające dłuższej refleksji niż „myślenie”. To nie zgadza się z powszechnym pojęciem automatycznych reakcji, które zwykle są szybsze niż refleksyjne, ale wynika to zapewne ze specyficznego wyboru słów i sensu pojęcia automatycznej genezy, którego nie każdy badany mógł w pełni zrozumieć. Z Rys. 1.2 można sądzić, że automatyczna geneza jest słabo skorelowana z ekstremalnymi wartościami znaku. Z Rys. 4.6 wynika też, że dla słów negatywnych i neutralnych mamy nieco wolniejsze czasy reakcji, podczas gdy dla pozytywnych krótsze. Te wyniki wymagają głębszego przemyślenia. Wymiar Genezy i Znak nie wydaje się wystarczający do wyjaśnienia obserwowanych zależności. Analiza skupień tych wyników pokazuje potrzebę zastosowania dodatkowych wymiarów do oceny emocji. Cemu wartości ocen na tych wykresach podane są w procentach, a nie w skali 1-3?

Na Rys. 4.7 przedstawiono przebiegi ERP, jak i ich rekonstrukcję dla 4 klasycznych komponentów z podzbiorów 2 do 9 elektrod umieszczonych w obszarach dobranych na podstawie literatury. Rekonstrukcja udaje się bardzo dobrze, ale komponent P300 wygląda na taki sam jak P200.

Podrozdział 4.4.3 zawiera analizę przebiegu czasowego globalnej mocy pola w czterech regionach ROI (lewa-prawa strona, czołowe-ciemiennicze obszary). Do każdego ROI przypisano 3 elektrody. Przedstawiona została szczegółowa analiza interakcji czynników Znak i Genezy w interakcji z ROI w krótkich okienkach czasowych. Otrzymaliśmy w ten sposób ciekawy opis różnych efektów interakcji. W początkowym okienku 60-120 ms słowa automatyczne wywoływały większe ujemne wartości amplitudy, a w lewej-czołowej lokalizacji słowa negatywne wywoływały istotnie mniejszą ujemną wartość amplitudy. Są tu liczne interesujące zależności, ale nie mamy dobrego modelu, który by je wyjaśniał.

Analiza skupień atomów dla 9 warunków eksperymentalnych nie pokazała globalnego maksimum dla indeksu Calińskiego-Harabasa oceniającego liczbę klastrów, wybrano więc

lokalne maksimum z 37 klastrami. Czy liczba 1-elementowych klastrów jest po prostu liczbą atomów, czy są to atomy całkiem niepodobne do innych? Jeśli mamy 9 warunków a 8 klastrów to jak dokładnie jeden atom na klaster może je wszystkie reprezentować? Na rysunku 4-14 i 4.16 widać, że istotnie dla każdego warunku mamy bardzo podobny atomy, ale klaster B na rys 4-15 ma wyraźnie dwa skupiska. Przechodząc do źródeł, dopasowano początkowe położenie dipoli (wzmianka o metodzie dopasowania za pomocą funkcji *ft\_dipolefitting* jest 25 stron wcześniej, co nie ułatwia czytania pracy). Lokalizacja tych dipoli na rysunku 4.18 mapowana na atlas Desikana wygląda sensownie, ale dopasowanie dla każdego z 9 warunków na atlasie Hagmana wygląda całkiem inaczej. W skupisku A widać aktywację bocznej kory ciemieniowej (LOC), w kolumnie obok kropki reprezentujące źródła dla poszczególnych warunków położone są głęboko, prawie w pniu mózgu; w kolumnie C widać dipol w środkowym zakręcie skroniowym (MST), ale na rysunku obok również zaznaczono jego położenia bliskie pnia, daleko od MST. Analiza poszczególnych klastrów pokazuje dobre dopasowanie, wskaźniki wariancji rezydualnej są poniżej 1%. Oscylacje w skupisku A są nieco poniżej 6 Hz, co zgadza się z wykresami przebiegów czasowych odpowiednich atomów.

Efekty Znaku, Genezy i ich interakcji na poziomie atomów są widoczne nieco lepiej i pozwalają się odnieść do wyników behawioralnych, np. słowa negatywne odpowiadają słabszym dodatnim amplitudom niż pozytywnie, co może wydłużać czas reakcji. Co jednak tłumaczy słabszą reakcję na negatywne słowa, a silniejszą na słowa o genezie refleksyjnej? Różnice dla poszczególnych warunków na wykresach 4.14 i 4.18 wydają się małe, a różnice amplitud na Rys. 4.19 są statycznie znaczące. Dla skupiska B dipole są w lewym zakręcie nadbrzeżnym (SMG) i lewej dolnej korze ciemieniowej (IPC). Jak pokazano w licznych eksperymentach, te regiony są zaangażowane w wizualne rozpoznawanie słów i i semantyczne oraz fonologiczne analizy tekstu. Jednakże, jak wspomniałem wcześniej, traktowanie skupiska B jako jednorodnego może być nadmiernym uproszczeniem. Pomimo tego wskaźnik dopasowania rezydualnej wariancji jest na wysokim poziomie  $2 \pm 0.8\%$ . W klastrze C, dla którego przebiegi dla wszystkich warunków są podobne do klasycznego komponentu N400, dopasowanie dwóch dipoli w środkowym zakręcie skroniowym (MTG) jest na poziomie  $9.1 \pm 1.8\%$ . Im późniejszy komponent tym bardziej sygnały będą zróżnicowane. W klastrze D mamy tylko jeden dipol, dopasowanie wypada najgłębiej, jest na poziomie  $11.8 \pm 8.1\%$ . Ponieważ ten sygnał pojawia się 500 ms po bodźcu, jest on odbiciem wielu procesów, w tym aktywności pamięci skojarzeniowej, która zachodzi w różnych obszarach mózgu i dla każdego warunku wyglądają nieco

inaczej. Trudno jest więc interpretować ten komponent jako efekt jednoznacznie umiejscowionych źródeł, co Autorka słusznie zaznacza.

Na zakończenie tego rozdziału przedstawiono porównania dekompozycji pierwszego z warunków eksperymentu z metodami PCA i ICA (tu wkradła się literówka, ICA JADE a nie jader; ten sam błąd jest w publikacji). Komentarz do tego rysunku pojawia się dopiero w następnym rozdziale, który zawiera podsumowanie wyników. Najpierw jednak zostały przypomniane efekty Znak i Genezy dla poszczególnych klastrów. Porównano też efekty dekompozycji MMP z analizą globalnej mocy pola (GFP), odpowiedzialnej za mikrostrany. O ile początkowy fragment przebiegu GFP przypomina atom A, pozostałe nie mają podobnego charakteru. Jednakże lokalne maksima atomów skupisk B, C i D wypadają w bardzo podobnym czasie co zmiany GFP. Nie wiem na czym oparte jest przekonanie, że mogą być takie mikrostrany, „które zawierają tylko pojedynczy komponent ukryty”. Być może dotyczy to pierwszych 200 milisekund, ale nie późniejszych komponentów.

Wracając do porównania MMP z PCA i ICA, niektóre komponenty są istotnie podobne, ale widać tu zalety MMP. PCA tworzy skomplikowane oscylacje, a ICA pokazuje pojedyncze piki z niewielkim szumem. Jednakże ICA ma wiele wariantów, określanych jako spacial, temporal i spatio-temporal ICA, które mogą dostarczyć bardziej czytelne komponenty, więc wymaga to szerszego porównania. Nie ma jednak wątpliwości, że metoda MMP jest jedną z najbardziej interesujących metod pozwalających na analizę potencjałów wywołanych. Autorka dobrze sobie zdaje sprawę z jej ograniczeń, związanych z interpretacją ukrytych komponentów. Jest więc jeszcze wiele do zrobienia.

Oprócz publikacji, na której oparta jest praca doktorska, mgr Duda-Goławska jest współautorką 9 innych prac naukowych zgodnych z ogólną tematyką przedstawioną w jej rozprawie, opublikowanych w czasopismach o wysokiej randze. We wszystkich przypadkach określiła swój wkład, który polegał głównie na analizie danych eksperymentalnych, analizach statystycznych oraz przygotowania publikacji, a w niektórych przypadkach również prowadzeniu badań EEG.

Można mieć wątpliwości czy sama idea identyfikacji ukrytych komponentów pozwalających na redukcję złożonych procesów w sieciach neuronowych mózgu do kombinacji reprezentacji pojedynczych procesów mentalnych jest możliwa. Nie zwalnia to nas jednak od próby sprawdzenia, jak daleko możemy w tym kierunku zejść. Podsumowując, praca doktorska mgr Dudy-Goławskiej dotyczy aktualnej i ważnej tematyki analizy sygnałów EEG i jest sporym

krokiem w kierunku ich lepszego zrozumienia. Otrzymane wyniki są dobrze uzasadnione a wzięwszy pod uwagę popularność analiz potencjałów wywołanych mogą znaleźć liczne zastosowania. Wyniki te zostały już opublikowane w czasopiśmie *Neuroinformatics*. Dlatego uważam, że formalne i zwyczajowe wymagania dotyczące prac doktorskich zostały tu spełnione i z pełnym przekonaniem wnioskuję o dopuszczenie Autorki do dalszych etapów przewodu doktorskiego.



Prof. Włodzisław Duch,

Toruń, 26/10/2022