

## Filtrowanie informacji zawartych w sieciach korelacyjnych lub innych sieciach ważonych opisujących układy złożone

Zgodnie z prawem Moore'a postęp technologiczny jest wykładniczy. Oznacza to, że moc obliczeniowa komputerów zwiększa się dwukrotnie w stałych odcinkach czasu. Jak długo trzeba czekać na wspomniane podwojenie? Szacuje się, że trwa to zaledwie od półtora do dwóch lat. Już teraz moc obliczeniowa najpotężniejszego komputera na świecie wynosi prawie 34 petaflopsy, co jest równoważne możliwości wykonania miliona miliardów operacji liczbowych w przeciągu sekundy. Te szokujące liczby zmieniają perspektywę poznawczą ludzi i pozwalają na prowadzenie coraz to bardziej skomplikowanych obliczeń. To co wcześniej wydawało się niemożliwe, teraz jest w zasięgu ręki naukowców.

Dzięki temu niezwyklejmu rozwojowi technologicznemu oczy wielu badaczy zwróciły się w kierunku najbardziej skomplikowanych obserwowanych układów. Przykładem jest ludzki mózg, który wciąż wyprzedza swoimi możliwościami nawet najlepsze stworzone przez ludzi maszyny. Szacuje się jednak, że superkomputery osiągną zdolności obliczeniowe ludzkiego mózgu w przeciągu najbliższych 10-15 lat. Mimo to, nikt nie ma zamiaru czekać bezczynnie i już teraz prowadzone są badania na temat funkcjonowaniu organu, który zapewnił naszemu gatunkowi naczelną rolę na planecie ziemi.

Z drugiej strony, rozwój technologii to również wzrost ilości zbieranych i przechowywanych danych. Przepływ informacji na świecie jest ogromny. W internecie jest około miliarda stron internetowych, a liczbę użytkowników szacuje się na ponad 3 miliardy. Z niemal dowolnego miejsca na świecie jesteśmy sprawdzić notowania giełdowe spółek z drugiego końca świata lub przeczytać jeden z prawie 5 milionów artykułów dostępnych na angielskiej wikipedii. A przecież oprócz danych dostępnych całej społeczności internetowej istnieją jeszcze olbrzymie bazy danych, należące do licznych firm i instytucji. To tłumaczy obserwowany w ostatnim czasie wzrost zainteresowania umiejętnościami związanymi z komputerową obróbką bardzo dużych ilości danych. Tak zwany *Data Scientist* (naukowiec od danych) jest obecnie najbardziej poszukiwanym stanowiskiem na rynku pracy.

Obydwie opisane powyżej perspektywy łączy jedna rzecz, chęć uproszczenia i zrozumienia problemu, którego poziom skomplikowania jest zbyt duży by móc spojrzeć na niego całościowo. Wiąże się to z problematyką tak zwanej *złożoności* (ang. *complexity*) o której wspominał niedawno Stephen Hawking, wskazując na nią jako na największe wyzwanie stojące przed współczesnymi badaczami. Takie podejście leży również u podstaw proponowanego przeze mnie projektu. **Chciałbym bowiem zaproponować i rozwijać nowy algorytm, który pozwoli na uproszczenie skomplikowanych struktur opisanych z użyciem sieci. Dlaczego właśnie sieci? Otóż podejście sieciowe okazało się w ostatnim czasie najskuteczniejsze w opisie i analizie układów złożonych.** Ilość prac naukowych poświęconych tak zwanym sieciom złożonym rośnie w ostatnim czasie nawet dynamiczniej niż wspomniana powyżej moc obliczeniowa najszybszych komputerów. Dobrze obrazuje to sytuacja z klasyczną pracą dotyczącą sieci losowych, napisaną przez znakomitego matematyka Paula Erdosa. Przed 2000 rokiem, który był przełomowy w dziedzinie sieci i ich zastosowań, cytowano ją średnio kilka lub kilkanaście razy rocznie. Jednakże w samym tylko 2005 roku zacytowano ją ponad 150 razy. Od tego czasu, sieci w różnej postaci, zostały użyte w analizie najróżniejszych układów i problemów poświęconych niezwykle różnorodnej tematyce. Poczynając od wspomnianego wcześniej mózgu, poprzez fizykę statystyczną, infrastrukturę telekomunikacyjną, transport, rozprzestrzenianie się wirusów, walkę z terroryzmem, a nawet rynki finansowe.

**Oprócz rozwoju wspomnianego algorytmu, projekt zakłada również wykorzystanie go do dostępnych danych. Korzystając z życzliwości kolegów z Wydziału Fizyki Uniwersytetu Warszawskiego, algorytm posłuży do przeanalizowania danych EEG, danych sejsmicznych oraz notowań finansowych. Tylko podejście interdyscyplinarne może, w mojej opinii, sprawdzić skuteczność zaproponowanego algorytmu.**