EFEKTYWNE WYKORZYSTANIE RANDOMIZACJI: OD SZEREGOWANIA DO ADWORDS

Jednym z ważnych zagadnień w algorytmice są *problemy przypisania*, gdzie obiekty łączone są w pary lub większe zbiory zgodnie z określonymi regułami. Kanonicznym przykładem są SKOJARZENIA W GRAFACH DWUDZIELNYCH, gdzie obiekty dwóch typów (np. studenci) są łączone w tak wiele rozdzielnych par (np. do zadań grupowych) jak to możliwe, zważywszy na ograniczenia wyrażone w języku grafów.

Teoria *szeregowania zadań* zajmuje się tymi z nich, gdzie przypisuje się *zadania* do *jednostek czasu* na dostępnych maszynach, celem wykonania. Problemy szeregowania zadań, badane od samego początku teorii optymalizacji w latach 40-tych, mają znaczenie tak w teorii jak i praktyce. Ich zastosowania obejmują planowanie produkcji i dystrybucji dóbr, optymalizację procesów, układanie harmonogramów, oraz wiele innych.

Wszystkie problemy przypisań (skojarzeń i szeregowania) były też intensywnie badane w modelu *obliczeń online*, istotnym dla zastosowań takich jak równoważenie obciążenia oraz trasowania w sieci, zarządzanie równoległymi procesami czy obliczenia w czasie rzeczywistym. W modelu online wejście pojawia sekwencyjnie (np. jedno zadanie po drugim), zaś *algorytm online* po każdym elemencie podejmuje ostateczną decyzję dot. przypisania zanim pozna kolejny element. Celem jest uzyskanie algorytmów z gwarancją (wyrażoną jako *współczynnik konkurencyjności*), że ich rozwiązania zawsze są bliskie optymalnym dla całego wejścia.

Problemy optymalizacji online są ściśle związane z teorią gier. Skoro algorytm dokonuje decyzji na bieżąco, ale jakość jego rozwiązania jest oceniana w najgorszym przypadku, możemy myśleć o drugim graczu, zwykle nazywanym *adwersarzem*, który dobiera kolejne elementy wejścia w taki sposób, by dotychczasowe decyzje algorytmu okazały się błędne i brzemienne w skutkach.

Jednym z fundamentalnych wyników dla problemów przypisań online jest uzyskanie najlepszych możliwych algorytmów dla problemu SKOJARZEŃ W GRAFACH DWUDZIELNYCH ONLINE, w tym zrandomizowanego o współczynniku $\frac{e}{e-1}$. (W tym problemie algorytm dostaje z góry wierzchołki z "partycji offline", które ma połączyć w jak najwięcej par z pojawiającymi się kolejno wierzchołkami z "partycji oline", których możliwe połączenia są ujawniane dopiero w chwili przybycia.)

W odróżnieniu od SKOJARZEŃ W GRAFACH DWUDZIELNYCH ONLINE, dla wielu ważnych problemów w szeregowaniu optymalny współczynnik pozostaje nieznany. Do niedawna jednym z przykładów był problem SZEREGOWANIA PAKIETÓW ONLINE, gdzie pojawiającymi się elementami są zadania jednostkowej wielkości o całkowitoliczbowych *deadline'ach* oraz nieujemnych *wagach* oznaczających priorytet bądź zysk związany z ukończeniem zadania przed deadlinem. Celem w tym problemie jest wybranie jednego zadania dla każdego slotu czasu w sposób, który maksymalizuje sumę wag wykonanych zadań.

Długo wierzono, że optymalny współczynnik dla SZEREGOWANIA PAKIETÓW ONLINE to „złoty podział", tj. $\phi \approx 1.618$, co potwierdziła dopiero w roku 2019 grupa badawczej z udziałem kierownika tego projektu.

W tym projekcie chcemy zbadać efekt *randomizacji* dla kilku ważnych problemów szeregowania online związanych z SZEREGOWANIEM PAKIETÓW. Randomizacja stanowi narzędzie, które (zwłaszcza w modelu online) pozwala poprawić wyniki na dwa sposoby. Po pierwsze, algorytmy zrandomizowane często uzyskują lepsze gwarancje niż deterministyczne – choćby dla SZEREGOWANIEM PAKIETÓW ONLINE, choć są zbliżone problemy, gdzie dotąd się to nie udało. Naszym celem jest określenie możliwego zysku z użycia randomizacji.

Po drugie i równie ważne, randomizacja często pozwala na znaczące uproszczenia w konstrukcji i analizie algorytmów. Przykładowo, wspomniana analiza ϕ-konkurencyjnego algorytmu zajmuje ponad 40 stron, podczas gdy niemal od początku badań nad tym problemem znany był niezwykle prosty algorytm zrandomizowany o lepszym współczynniku $\frac{e}{e-1}$, który wraz z elegancką analizą zmieścić można na jednej stronie.

Algorytmy online dla problemów przypisań są również związane z *projektowaniem mechanizmów*, np. aukcji, co wiąże się z ich wspólnymi korzeniami w teorii gier oraz wspólnymi zastosowaniami w zarządzaniu ruchem w sieciach, choć dziedziny te różnią się założeniami nt. autonomii i zachowania uczestników ruchu.

Istotnie, zarówno SZEREGOWANIE PAKIETÓW jak i SKOJARZENIA ONLINE stanowią szczególny przypadek PROBLEMU ADWORDS, w którym krawędzie grafu mają wagi zaś wierzchołki z partycji offline mają własne ograniczenia na sumaryczną wagę wybranych przyległych krawędzi. Problem ten jest ważny z dwóch powodów. Stanowi on naturalne uogólnienie problemu skojarzeń. Ponadto, co istotne dla firm takich jak Google, Microsoft i Facebook, ADWORDS opisuje maksymalizację zysku z reklam wyświetlanych wraz z wynikami wyszukiwania bądź podczas przewijania ekranu – wierzchołkom offline odpowiadają reklamodawcy, którzy deklarują płacić określoną cenę za każde wyświetlenie ich reklamy (w odpowiednich warunkach), jednak nie więcej niż wynosi ich budżet. Dla tego problemu udało się uzyskać optymalny współczynnik $\frac{e}{e-1}$ tylko dla kilku szczególnych przypadków z dodatkowymi założeniami, ale w ogólności znane są tylko trywialne algorytmy 2-konkurencyjne.

Wierzymy, że doświadczenie z dotychczasowych i planowanych badań nad SZEREGOWANIEM PAKIETÓW, zaprocentuje w problemie ADWORDS, pozwalając uzyskać optymalny algorytm dla szerszej klasy instancji.