

Krzysztof Sornat

# Approximation Algorithms for Multiwinner Elections and Clustering Problems

*PhD thesis*

## Abstract

The thesis is devoted to polynomial time approximation algorithms for a few NP-hard discrete optimization problems that model real-world issues such as clustering and multiwinner elections.

Multiwinner elections and clustering problems have very similar settings. Our goal in both is to choose a fixed number of objects among a finite set of alternatives. We want to optimize the function of distances/disagreement between input points/voters and cluster centers/winners.

We exploit similarities and explore the differences in order to understand when efficient algorithms in one model can be useful in another. The main contributions of the thesis are the following algorithms and proofs of their approximation factors:

1. The first polynomial time constant-factor approximation algorithm for the ORDERED  $k$ -MEDIAN problem. The constant achieved is equal to  $38 + \epsilon$  for any  $\epsilon > 0$ . This improves the previous best  $\mathcal{O}(\log n)$ -approximation achieved by Aouad and Segev (2018).
2. The first polynomial time constant-factor approximation algorithm for the RECTANGULAR ORDERED  $k$ -MEDIAN problem. The constant achieved is equal to 15. The previous best approximation was  $\mathcal{O}(\log n)$  (Aouad and Segev 2018).
3. The first polynomial time constant-factor approximation for the HARMONIC  $k$ -MEDIAN problem in general spaces (not only metric). The constant is upperbounded by 2.36. To the best of our knowledge this is the first natural variant of  $k$ -MEDIAN that has constant-factor approximation without assuming the triangle inequality.
4. The first PTAS (polynomial time approximation scheme) for the MIN-IMAX APPROVAL VOTING problem. This improves the previous best 2-approximation (Caragiannis et al. 2010).

5. A parameterized approximation scheme for MINIMAX APPROVAL VOTING parameterized by the value  $d$  of an optimal solution. The running time is upperbounded by  $(3/\epsilon)^{2d}(nm)^{\mathcal{O}(1)}$ . It is essentially optimal assuming the Exponential Time Hypothesis due to Cygan et al. (2018). The parameterized approximation scheme allows us to construct a faster PTAS for MINIMAX APPROVAL VOTING.

We believe that the ideas and algorithm analysis techniques developed in this thesis will be useful in further work on approximation algorithms. We also hope our results will stimulate more interdisciplinary research on relations between clustering problems and multiwinner elections.

**Keywords:** approximation algorithms, linear programming, LP-rounding, parameterized algorithms, clustering, k-median, facility location, multiwinner election, proportional approval voting, minimax approval voting, computational social choice

Krzysztof Sornat

# Algorytmy aproksymacyjne dla problemów wyboru wielu zwycięzców oraz klastrowania

*rozprawa doktorska*

## Streszczenie

Rozprawa doktorska jest poświęcona algorytmom aproksymacyjnym rozwiązującym wybrane NP-trudne problemy optymalizacji dyskretnej w czasie wielomianowym.

Rozważane problemy obliczeniowe modelują zagadnienia znane z przemysłowych i społecznych zastosowań: klastrowanie oraz wybór wielu zwycięzców. W obu rodzajach problemów naszym celem jest wybranie określonej liczby obiektów spośród skończonego zbioru alternatyw; przy tym wyborze kierujemy się optymalizowaniem pewnej funkcji odległości (niezadowolenia) pomiędzy punktami danych (głosującymi), a centrami klastrów (zwycięzcami).

W trakcie pracy nad nowymi algorytmami rozważaliśmy powyższe podobieństwa, jak i różnice pozwalające lepiej zrozumieć, kiedy algorytmy efektywne dla jednego zagadnienia mogą być użyteczne także w drugim. Głównym wkładem tej rozprawy doktorskiej jest konstrukcja poniżej wymienionych algorytmów oraz analiza ich współczynników aproksymacji:

1. Pierwszy algorytm o stałym współczynniku aproksymacji dla problemu ORDERED  $k$ -MEDIAN. Osiągnięta stała wynosi  $38 + \epsilon$  dla dowolnego  $\epsilon > 0$ . W ten sposób poprawiliśmy poprzedni najlepszy współczynnik aproksymacji, który wynosił  $\mathcal{O}(\log n)$ , a został osiągnięty przez Aouada oraz Segeva (2018).
2. Pierwszy algorytm o stałym współczynniku aproksymacji dla problemu RECTANGULAR ORDERED  $k$ -MEDIAN. Osiągnięta stała wynosi 15. Poprzedni najlepszy współczynnik aproksymacji wynosił  $\mathcal{O}(\log n)$  (Aouad oraz Segev 2018).
3. Pierwszy algorytm o stałym współczynniku aproksymacji dla problemu HARMONIC  $k$ -MEDIAN w ogólnych przestrzeniach (nie tylko metrycznych). Osiągnięta stała jest ograniczona z góry przez 2.36. Zgodnie z naszą wiedzą jest to pierwszy naturalny wariant problemu  $k$ -

MEDIAN, który potrafimy aproksymować ze stałym współczynnikiem bez zakładania nierówności trójkąta.

4. Pierwszy schemat aproksymacji wielomianowej (PTAS) dla problemu MINIMAX APPROVAL VOTING. Poprzedni najlepszy współczynnik aproksymacji osiągnięty przez Caragiannisa i współpracowników (2010) wynosił 2.
5. Schemat aproksymacji w czasie parametryzowanym dla problemu MINIMAX APPROVAL VOTING parametryzowanego wartością optymalnego rozwiązania, którą oznaczamy  $d$ . Czas działania jest ograniczony z góry przez  $(3/\epsilon)^{2d}(nm)^{O(1)}$ . Cygan i inni autorzy (2018) pokazali, że zasadniczo jest to najszybszy tego typu algorytm przy założeniu Hipotezy Czasu Wykładniczego. Skonstruowany algorytm pomógł nam w opracowaniu szybszego schematu aproksymacji wielomianowej dla MINIMAX APPROVAL VOTING.

Wierzmy, że użyte pomysły i techniki analizy algorytmów zawarte w tej rozprawie będą użyteczne w dalszych pracach nad algorytmami aproksymacyjnymi. Mamy również nadzieję, że nasze wyniki będą stymulowały następne badania nad relacjami pomiędzy problemami klastrowania oraz wyborami wielu zwycięzców.

**Słowa kluczowe:** algorytmy aproksymacyjne, programowanie liniowe, okrągłanie rozwiązań programów liniowych, algorytmy parametryzowane, klastrowanie, k-median, umiejscawianie fabryk, wybory wielu zwycięzców, proportional approval voting, minimax approval voting, obliczeniowa teoria wyboru społecznego