

Imię i nazwisko studenta: Paulina Brzęcka

Nr albumu: 184701

Poziom kształcenia: studia drugiego stopnia

Forma studiów: stacjonarne Kierunek studiów: Informatyka

Specjalność: Algorytmy i technologie internetowe

PRACA DYPLOMOWA MAGISTERSKA

Tytuł pracy w języku polskim: Analiza algorytmów dla dominowania rzymskiego słabo spójnego
Tytuł pracy w języku angielskim: Analysis of algorithms for weakly connected Roman domination

Opiekun pracy: dr inż. Joanna Raczek

STRESZCZENIE

Streszczenie pracy opisuje problem naukowy polegający na opracowaniu i wdrożeniu metody analizy danych w środowisku wielowymiarowym. Celem pracy jest zaproponowanie algorytmu optymalizacyjnego, który pozwala na efektywne przetwarzanie dużych zbiorów danych. Zakres pracy obejmuje analizę istniejących metod, implementację nowego rozwiązania oraz ocenę jego skuteczności na wybranych przypadkach testowych. Zastosowana metoda badawcza obejmowała modelowanie matematyczne, programowanie w języku Python oraz wizualizację wyników. Wyniki wskazują na istotne przyspieszenie obliczeń przy zachowaniu wysokiej dokładności. Najważniejszym wnioskiem jest możliwość zastosowania zaproponowanego algorytmu w rzeczywistych aplikacjach analitycznych.

Słowa kluczowe: algorytmy, przetwarzanie danych, optymalizacja.

ABSTRACT

The abstract describes a scientific problem focusing on the development and implementation of a data analysis method in a multidimensional environment. The objective of this thesis is to propose an optimization algorithm that enables efficient processing of large datasets. The scope of the work includes an analysis of existing methods, the implementation of a new solution, and the evaluation of its effectiveness on selected test cases. The research methodology involved mathematical modeling, Python programming, and visualization of results. The results indicate significant acceleration in computations while maintaining high accuracy. The key conclusion is the feasibility of applying the proposed algorithm in real-world analytical applications.

Keywords: algorithms, data processing, optimization.

SPIS TREŚCI

VV	уках	wazniejszych oznaczen i skrotow	5
1	Wst	tęp i cel pracy	6
	1.1	Cel pracy	6
	1.2	Zakres pracy	6
2	Wpi	rowadzenie teoretyczne	7
	2.1	Wprowadzenie	7
	2.2	Geneza historyczna	7
	2.3	Przegląd literatury	7
	2.4	Metody badań	7
3	Badane algorytmy		
	3.1	Wprowadzenie	g
	3.2	Algorytm Brute Force	õ
		3.2.1 Działanie	õ
		3.2.2 Złożoność i wydajność	õ
		3.2.3 Pseudokod	õ
	3.3	Algorytm liniowy dla drzew	õ
		3.3.1 Pseudokod	õ
	3.4	Algorytm programowania liniowego I	õ
		3.4.1 Pseudokod	õ
	3.5	Algorytm programowania liniowego II	g
		3.5.1 Pseudokod	g
	3.6	Algorytm mrówkowy	g
		3.6.1 Pseudokod	9
	3.7	Algorytm aproksymacyjny	g
		3.7.1 Pseudokod	g
4	Pod	Isumowanie i wnioski	16
	4.1	Podsumowanie wyników	16
	4.2	Wnioski i dalsze kierunki badań	16
Α	Zała	ączniki	18
	A.1	Dodatkowe materiały	18
		A.1.1 Schemat obliczeń	18
		A 1.2 Kod źródłowy	19

WYKAZ WAŻNIEJSZYCH OZNACZEŃ I SKRÓTÓW

 $e \;\; - \;\; {
m Niepewność}$ pomiaru.

f – Częstotliwość [Hz].

 $k - \mathrm{Stała}$ Boltzmanna $1.38 \cdot 10^{-23}$ Ws/K.

1. WSTĘP I CEL PRACY

Tematem pracy jest analiza algorytmów znajdujących funkcję dominującą rzymską słabospójną w grafach. Problem znajdywania liczby dominowania rzymskiego słabospójnego jest problemem NP-trudnym. Wersja decyzyjna tego problemu jest NP-zupełna. Nie istnieją zatem dokładne algorytmy rozwiązujące problem w czasie wielomianowym. Dlatego niniejsza praca dokonuje analizy dostępnych i proponowanych algorytmów rozwiązujących ten problem w sposób zarówno dokładny, jak i przybliżony, w celu znalezienia skutecznych rozwiązań oraz zastosowań.

1.1 Cel pracy

Celem pracy jest analiza algorytmów dla dominowania rzymskiego słabo spójnego, w tym opisanie już istniejących rozwiązań oraz opracowanie własnych, porównanie ich skuteczności oraz możliwych praktycznych zastosowań.

1.2 Zakres pracy

W ramach pracy dokonano systematycznego przeglądu literatury. W literaturze proponowano wiele algorytmów dokładnych o czasie wykładniczym, między innymi algorytmy wykorzystujace programowanie liniowe. Dodatkowo, w wielu publikacjach skupiono się na algorytmach dla konkretnych klas grafów. W literaturze zostały również zdefiniowane algorytmy niedokładne, aproksymacyjne, o różnej jakości rozwiązania. Na podstawie znalezionej literatury zaimplementowane zostały dwa algorytmy programowania liniowego oraz aproksymacyjny o współczynniku aproksymacji $2(1+ln(\Delta+1))$. W ramach własnej pracy, zaimplementowano algorytm dokładny brute force, liniowy dokładny dla drzew oraz mrówkowy. Niniejsza praca opisuje wymienione algorytmy, porównuje je pod kątem wydajności, poprawności oraz czasu działania.

2. WPROWADZENIE TEORETYCZNE

2.1 Wprowadzenie

Jako, że problem nie jest powszechnie znany, należy wprowadzić następujące pojęcia [1]: **Definicja 1** Funkcja dominująca rzymska zdefiniowana jest dla grafu G=(V,E), gdzie $f:V->\{0,1,2\}$ spełnia warunek, że dla każdego wierzchołka u, dla którego f(u)=0 jest sąsiadem przynajmniej jednego wierzchołka v, dla którego f(v)=2.

Definicja 2 Dominujący zbiór $D \subseteq V$ jest zbiorem dominującym słabospójnym grafu G jeśli graf $(V, E \cap (D \times V))$ jest spójny.

Definicja 3 Funkcja dominująca rzymska słabospójna na grafie G będzie funkcją dominującą rzymską, taką, że zbiór $\{u \in V : f(u) \in \{1,2\}\}$ jest jednocześnie zbiorem dominującym słabospójnym.

Definicja 4 Wagę funkcji dominującej rzymskiej słabospójnej definiujemy jako $f(V) = \sum_{u \in V} f(u)$. Minimalną wartość tej funkcji nazywamy liczbą dominowania rzymskiego słabospójnego.

2.2 Geneza historyczna

Problem swoją nazwę zawdzięcza imperium rzymskiemu. Po raz pierwszy został opisany w artykule "Defend Roman Empire!". Obrazuje on problem następująco: Każdy wierzchołek grafu reprezentuje pewną lokalizację (miasto, wzgórze) w Imperium Rzymskim. Lokalizacja (wierzchołek v) jest niechroniona, jeśli nie stacjonują w niej żadne legiony wojska (f(v)=0) oraz chroniona jeśli ($f(v)\in 1,2$). Wartości te oznaczają liczbę legionów stacjonujących w danej lokalizacji. Niechroniona lokalizacja może być ochroniona poprzez wysłanie legionu stacjonującego w lokalizacji sąsiadującej. W czwartym wieku cesarz Konstantyn Wielki wydał dekret zakazujący przemieszczenia się legionu do lokalizacji sąsiadującej, jeśli sprawi to, że aktualna lokalizacja pozostanie niechroniona. Dlatego, żeby móc wysłać legion do sąsiedniej lokacji, w aktualnej muszą stacjonować dwa legiony. Oczywiście, cesarzowi zależało na jak najmniejszych kosztach utrzymywania legionów, a zatem, żeby było ich jak najmniej. [1]

2.3 Przegląd literatury

Omówienie istniejących metod i podejść z literatury, które odnoszą się do problemu badawczego. Jednym z istotnych źródeł w tym zakresie jest strona National Center of Biotechnology Information, która dostarcza bogatych danych do analiz [2].

2.4 Metody badań

W celu implementacji i testowania wydajności oraz poprawności algorytmów, stworzono program w języku Python, ze wsparciem następujących bibliotek:

- networkx pakiet dostarczający funkcje umożliwiające operacje na grafach, wykresach i sieciach
- matplotlib do wyświetlania wyników działania algorytmów w postaci wykresów grafów
- time wykorzystywane do pomiarów czasu pracy algorytmów
- pulp do programowania liniowego
- gurobipy do programowania liniowego

Program umożliwia wprowadzenie dowolnego grafu w postaci listy wierzchołków oraz krawędzi, jak i wygenerowanie losowego grafu. Następnie wybrane algorytmy analizują dany graf poprzez przypisywanie odpowiednich wartości wierzchołkom oraz wyliczania liczby dominowania rzymskiego słabospójnego. Dla każdego z algorytmów wyliczany i zapisywany jest ich czas działania. Na końcu program wyświetla wykres z nadanymi wartościami na wierzchołkach.

3. BADANE ALGORYTMY

3.1 Wprowadzenie

Niniejszy rozdział opisuje algorytmy dla funkcji dominowania rzymskiego słabospójnego. Analizowane one będą pod kątem złożoności, wydajności, poprawności oraz potencjalnego zastosowania. Wszystkie algorytmy wyznaczają dodatkowo zbiór dominowania rzymskiego słabospójnego. Lista analizowanych algorytmów jest następująca:

- · algorytm brute force
- · algorytm liniowy dla drzew
- · algorytm programowania liniowego I
- · algorytm programowania liniowego II
- · algorytm mrówkowy
- · algorytm aproksymacyjny

3.2 Algorytm Brute Force

3.2.1 Działanie

Jest to w zasadzie trywialna implementacja dokładnego algorytmu wyznaczającego funkcję oraz zbiór

- 3.2.2 Złożoność i wydajność
- 3.2.3 Pseudokod
- 3.3 Algorytm liniowy dla drzew
- 3.3.1 Pseudokod
- 3.4 Algorytm programowania liniowego I
- 3.4.1 Pseudokod
- 3.5 Algorytm programowania liniowego II
- 3.5.1 Pseudokod
- 3.6 Algorytm mrówkowy
- 3.6.1 Pseudokod
- 3.7 Algorytm aproksymacyjny
- 3.7.1 Pseudokod

Algorytm Brute Force

```
1: function FINDROMANDOMINATINGSET(graph)
        Initialize min\_roman\_number \leftarrow \infty
2:
3:
        \textbf{Initialize} \ best\_node\_values \leftarrow None
        nodes \leftarrow list of nodes in <math>graph
 4:
        for each assignment of values (0, 1, 2) to all nodes do
 5:
            node\_values \leftarrow mapping of nodes to values
6:
            induced\_set \leftarrow \mathsf{nodes} \ \mathsf{with} \ \mathsf{values} \ \{1,2\}
7:
8:
            for each node in induced\_set do
9:
                Add all its neighbors to induced\_set
            end for
10:
            Create empty induced\_graph
11:
            for each node in induced\_set do
12:
13:
                if node\_values[node] is 1 or 2 then
                    {f for} each neighbor in graph {f do}
14:
                        if neighbor in induced\_set then
15:
                           Add edge to induced\_graph
16:
17:
                        end if
                    end for
18:
                end if
19:
            end for
20:
            if induced\_graph is connected then
21:
                \textbf{Compute} \ roman\_number \leftarrow \textbf{sum of} \ node\_values
22:
                if roman\_number < min\_roman\_number then
23:
24:
                    Update min\_roman\_number and best\_node\_values
                end if
25:
            end if
26:
        end for
27:
        \textbf{return} \; (min\_roman\_number, best\_node\_values)
28:
29: end function
```

Algorytm liniowy dla drzew - Faza 1

```
1: function Phase1(T, root)
       father map ← Compute parent-child relationships using BFS
2:
       nodes\_ids \leftarrow \text{List of all nodes in } T
3:
4:
       for each node v in reversed(nodes ids) do
           father \leftarrow father\_map[v]
 5:
           if v is a leaf and v \neq root then
 6:
               if father exists then
 7:
                  Increase T[father]['n00']
8:
                  Set T[father]['child'] \leftarrow v
9:
               end if
10:
           else
11:
               if T[v]['n00'] == 1 and T[v]['n01'] == 0 and father exists then
12:
                  Increase T[father]['sw']
13:
               end if
14:
15:
               if T[v]['sw'] + T[v]['n00'] + T[v]['n01'] > 1 then
                  Set T[v]['R'] = 2
16.
                  if father exists then
17:
                      Increase T[father]['n2']
18:
                      if T[v]['n00'] == 1 and T[v]['n01'] == 0 then
19
                          Decrease T[father]['sw']
20:
21:
                      end if
                  end if
22:
                  T[v]['ch'] = 1
23:
               end if
24:
25:
               if T[v]['sw'] == 0 then
                  if T[v]['n00'] > 1 or (T[v]['n00'] == 1 and T[v]['n2'] == 0) then
26:
                      Set T[v]['R'] = 2
27:
                      if father exists then
28:
                          Increase T[father]['n2']
29:
30:
                      end if
                  else if T[v]['n00'] == 1 then
31:
                      Set T[v]['R'] = 0
32:
                      Set T[T[v]['child']]['R'] = 1
33:
                      if father exists then
34:
                          Decrease T[father]['sw']
35.
                      end if
36:
                  end if
37:
               end if
38:
           end if
39.
40:
       end for
41:
       if T[root]['n2'] == 0 and T[root]['R'] == 0 then
           Set T[root]['R'] = 2
42:
       end if
43:
       if T[root]['n1'] == (number of root's neighbors - 1) then
44:
45:
           Set T[root]['R'] = 1
       end if
46:
47:
       return T
48: end function
```

Algorytm liniowy dla drzew - Faza 2

```
1: function Phase2(T, root)
       father\_map \leftarrow Compute parent-child relationships using BFS
       nodes ids \leftarrow List of all nodes in T
3:
       for each node v in reversed(nodes\_ids) do
4:
           father \leftarrow father\_map[v]
 5:
           if father exists then
 6:
7:
               T[father]['child'] \leftarrow v
               if T[v]['n00'] == 1 and T[father]['ch'] == 1 and T[v]['n01'] == 0 then
8.
                   Set T[v]['R'] = 0
9:
                   Set T[T[v]['child']]['R'] = 1
10:
                   Increase T[father]['n00']
11:
               end if
12:
           end if
13:
14:
       end for
       return T
15:
16: end function
```

Algorytm programowania liniowego I

```
1: function ILP_I(graph)
        V \leftarrow \text{list of nodes in } graph
 3:
        E \leftarrow \text{list of edges in } graph
        Initialize ILP model model
 4:
        Set objective: Minimize \sum (a[i] + b[i]) for all nodes i \in V
 5:
 6:
        Define binary variables:
        x[i,j] for (i,j) \in E
                                                                                           \triangleright 1 if edge is in G'
 7:
                                                                           \triangleright 1 if edge is in spanning tree T'
        y[i,j] for (i,j) \in E
8:
        a[i] for i \in V
                                                                             \triangleright 1 if node belongs to V1 \cup V2
9:
        b[i] for i \in V
                                                                                   \triangleright 1 if node belongs to V2
10:
        Constraints:
11:
        for each node i in V do
                                                                            ⊳ Each node must be defended
12.
            Add constraint: a[i] + \sum b[k] \ge 1, where k are neighbors of i
13:
        end for
14:
        for each edge (i, j) in E do
15:
            Add constraint: y[i, j] \le x[i, j]
                                                                               \triangleright Tree edge must exist in G'
16:
            Add constraint: x[i, j] \le a[i] + a[j]
                                                             > Tree edges must connect defended nodes
17:
        end for
18:
        Add constraint: \sum y[i,j] = |V| - 1
                                                                           \triangleright Tree must have |V|-1 edges
19:
        Find cliques of size \geq 3 in graph and store as subsets
20:
21:
        for each subset S in subsets do
                                                                                          Add constraint: \sum y[i,j] \le |S| - 1 for edges (i,j) \in S
22:
        end for
23:
        for each node i in V do
                                                                            \triangleright V2 nodes must be in V1 \cup V2
24:
            Add constraint: b[i] \leq a[i]
25:
        end for
26:
        Solve ILP model
27:
        Extract solution:
28:
        for each node i in V do
29:
            solution[i] \leftarrow round(a[i].X) + 2 * round(b[i].X)
30:
31:
        end for
32.
        return (model.objVal, solution)
33: end function
```

Algorytm programowania liniowego II

```
1: function ILP_II(graph)
        Initialize ILP model model with minimization objective
 2:
        V \leftarrow \text{list of nodes in } graph
 3:
        E \leftarrow \text{list of edges in } graph
 4:
                                                                                            > Number of nodes
 5:
        n \leftarrow |V|
        Define binary variables:
 6:
        x[i] for i \in V
                                                                                        \triangleright 1 if node i is in set X
 7:
        y[i] for i \in V
                                                                                        \triangleright 1 if node i is in set Y
 8:
        a[e] for e \in E
                                                                          \triangleright 1 if edge e is in the spanning tree
 9:
10:
        t[i] for i \in V
                                                                                         \triangleright 1 if node i is the root
        Define integer and continuous variables:
11:
        u[i] for i \in V
                                                                          ▷ Integer variable for tree structure
12:
        v[e] for e \in E
                                                                         \triangleright Flow variable with bounds [-n, n]
13:
        Objective:
14.
        Minimize \sum (x[i] + y[i]) for all i \in V
15:
        Constraints:
16:
        for each node i in V do
                                                                              > Ensure all nodes are covered
17:
            Add constraint: x[i] + \sum y[j] \ge 1, where (i, j) \in E
18:
            Add constraint: y[i] \le x[i]
19:
20:
            Add constraint: \sum a[e] \ge 1, where e contains i
        end for
21:
        for each edge e = (i_e, j_e) in E do
22:
            Add constraint: a[e] \le x[i_e] + x[j_e]
23:
            Add constraint: v[e] \leq n \cdot a[e]
24:
25:
            Add constraint: v[e] \ge -n \cdot a[e]
        end for
26:
        Add constraint: \sum t[i] = 1
                                                                                          > Only one root exists
27:
        for each node i in V do
                                                                                   > Tree structure constraints
28:
            Add constraint: u[i] \leq n \cdot t[i]
29:
            Add constraint: u[i] + \sum v[e] - \sum v[e] = 1, for edges e entering/exiting i
30:
        end for
31:
        Solve ILP model
32:
        Extract solution:
33:
        for each node i in V do
34:
35:
            solution[i] \leftarrow round(x[i].varValue) + 2 \times round(y[i].varValue)
36:
        return (model.objVal, solution)
37:
38: end function
```

Algorytm mrówkowy - inicjalizacja

```
1: function InitializePheromones(graph)
       pheromones ← Assign initial pheromone value to all edges
       return pheromones
4: end function
 5: function CHOOSENODEVALUE(node, pheromones, neighbors)
6:
       values \leftarrow \{0, 1, 2\}
7:
       Initialize probabilities as empty list
       for each value in \{0, 1, 2\} do
 8:
           Compute pheromone_level as sum of pheromones of neighboring edges
 9:
           Compute heuristic based on number of neighbors
10:
           Compute probability as (pheromone\_level^{\alpha}) \times (heuristic^{\beta})
11:
           Append probability to probabilities
12:
       end for
13:
       Normalize probabilities
14:
       return Random weighted choice from \{0, 1, 2\}
15:
16: end function
17: function BuildSolution(graph, pheromones)
       Initialize node values as empty dictionary
18:
       for each node in graph do
19:
           neighbors \leftarrow list of node's neighbors
20:
21:
           Assign node\_values[node] \leftarrow \mathsf{CHOOSENODEVALUE}(\mathsf{node}, \mathsf{pheromones}, \mathsf{neighbors})
       end for
22:
       {\bf return} \ node\_values
23:
24: end function
25: function EVALUATESOLUTION(graph, node_values)
       if not IsValidRomanDominatingSet(graph, node_values) then
26:
           return \infty
27:
       end if
28:
       return Sum of all node values
30: end function
```

Algorytm mrówkowy - główna petla

```
1: function UPDATEPHEROMONES(graph, pheromones, solutions)
       for each edge in pheromones do
 2:
          Reduce pheromone level using evaporation rate
3:
       end for
 4:
       best \ solution \leftarrow Solution with minimum Roman number
 5:
       for each node in best solution do
 6:
 7.
          for each neighbor of node do
              Increase pheromone level on edge (node, neighbor)
 8:
 9:
          end for
       end for
10:
11: end function
12: function EXECUTE(graph)
       pheromones \leftarrow InitializePheromones(graph)
13:
       best \ solution \leftarrow None
14:
       best \ roman \ number \leftarrow \infty
15:
       for each iteration in num iterations do
16:
17.
          Initialize solutions as empty list
          for each ant in num ants do
18
              solution \leftarrow BuildSolution(graph, pheromones)
19:
              roman\_number \leftarrow EvaluateSolution(graph, solution)
20:
              Append (solution, roman number) to solutions
21:
22.
              if roman\_number < best\_roman\_number then
                 Update best_roman_number and best_solution
23:
24:
              end if
          end for
25:
          UPDATEPHEROMONES(graph, pheromones, solutions)
26:
27:
       return (best roman number, best solution)
28:
29: end function
```

Algorithm 1 Algorytm aproksymacyjny

```
1: function ComputeDominatingSet(graph)
2:
        dominating\_set \leftarrow \emptyset
       uncovered\_nodes \leftarrow \texttt{all nodes in } graph
3:
       while uncovered\_nodes is not empty do
4:
           max\_degree\_node \leftarrow node with highest degree in uncovered\_nodes
5:
           Add max degree node to dominating set
6:
           Remove max degree node and its neighbors from uncovered nodes
7:
        end while
8:
       return\ dominating\_set
9:
10: end function
11: function EXECUTE(graph)
       dominating\_set \leftarrow \texttt{ComputeDominatingSet}(graph)
12:
        node\ values \leftarrow \{node: 2\ \text{if}\ node \in dominating\ set,\ \text{else}\ 0\}
13:
        roman \ number \leftarrow sum \ of \ values \ in \ node \ values
14:
        return (roman\_number, node\_values)
15
16: end function
```

4. PODSUMOWANIE I WNIOSKI

4.1 Podsumowanie wyników

Przedstawienie kluczowych wyników pracy oraz ich znaczenia w kontekście postawionego celu badawczego.

4.2 Wnioski i dalsze kierunki badań

Na podstawie wyników pracy sformułowano następujące wnioski:

- Wniosek 1: [Opis pierwszego wniosku].
- Wniosek 2: [Opis drugiego wniosku].

Dalsze badania mogłyby obejmować:

- Rozszerzenie algorytmu na inne typy danych.
- · Testy w środowisku rzeczywistym.

WYKAZ LITERATURY

- 1. DR INŻ. JOANNA RACZEK, DR JOANNA CYMAN. Weakly connected Roman domination in graphs.

 Dostępne także z: https://mostwiedzy.pl/en/publication/weakly-connected-roman-domination-in-graphs, 150016-1. Dostęp: 03.03.2025.
- 2. OPENAI. Treść wygenerowana przy użyciu narzędzi ChatGPT (wersja 2024). 2024. Dostępne także z: https://openai.com. Dostęp: 15.01.2025.
- KOWALSKI, J.; KABACKI, J. Simulation of Network Systems in Education. W: Proceedings of the XXIV Autumn International Colloquium Advanced Simulation of Systems. Ostrava, Czechy: ASIS, 2002, s. 213–218. 9–11 września 2002.

A. ZAŁĄCZNIKI

A.1 Dodatkowe materialy

Przykładowe materiały pomocnicze:

- · Schematy obliczeniowe,
- · Dodatkowe wykresy wyników,
- Fragmenty kodu źródłowego (jeśli dotyczy).

A.1.1 Schemat obliczeń

Prezentacja dodatkowych szczegółów dotyczących analizy obliczeniowej.

A.1.2 Kod źródłowy

Wybrane fragmenty implementacji algorytmów w języku Python:

```
def example_function(data):
```

```
return [x**2 for x in data if x > 0]
```