

# **SPRAWOZDANIE**

Algorytmy i Struktury Danych

## ***Algorytmy grafowe***

Bartłomiej Kowalewski, 145204, I5.2

Jakub Koza, 145436, I5.2

*Pomiary wykonane zostały na sprzęcie:*

Laptop Lenovo G710

Windows 10 Home 64-bit

Procesor: Intel Core i7-4702MQ (2.2-3.2 GHz)

Pamięć RAM: 16 GB

Dysk twardy: 1TB SSHD

Karta graficzna: Intel HD Graphics 4600

*Algorytmy zaimplementowane zostały w języku Python 3.7:*

*W całości sprawozdania czas podawany jest w mikrosekundach.*

*Na wykresach pominęliśmy zaznaczanie odchylenia standardowego, zamieściliśmy natomiast tabele zawierające dokładne zestawienie pomiarów.*

*Pomiary czasu wykonane zostały za pomocą biblioteki datetime.*

1. Wykonane przez nas pomiary w pięciu próbach dla trzech reprezentacji postaci maszynowej grafu skierowanego i dwóch metod sortowania topologicznego grafu ilustrują następujące tabele:

a) *Macierz sąsiedztwa*

MACIERZ SĄSIEDZTWA ALG. KAHNA							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	0	0	0	0	0	0	0
40	15656	0	0	15613	0	6254	7659
50	0	0	0	15647	0	3129	6259
60	15622	15624	0	15593	0	9368	7649
70	15610	15628	15588	15620	15630	15615	15
80	31222	31248	15619	15622	15622	21867	7649
90	31218	31279	31246	31224	31216	31237	24
100	46839	62489	31279	31240	31217	40613	12495
110	46868	46873	46841	46901	46848	46866	21
120	46872	78123	62527	62488	46872	59376	11694
130	78117	93768	78092	78150	78120	81249	6259
140	109335	78150	78149	62526	78121	81256	15286
150	156237	93768	93727	93720	78153	103121	27235

MACIERZ SĄSIEDZTWA DFS							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	0	0	15620	0	0	3124	6248
40	0	0	0	0	0	0	0
50	0	0	0	0	15614	3123	6246
60	15601	0	0	0	0	3120	6240
70	0	15614	0	0	0	3123	6246
80	15624	15630	15633	15622	15614	15625	7
90	15626	0	0	0	15624	6250	7655
100	31271	15622	15626	0	15628	15629	9889
110	31226	15626	15621	31253	15633	21872	7649
120	15622	15632	15623	15650	15611	15628	13
130	15621	31274	15624	15619	31250	21878	7662
140	46874	31218	31253	15633	31228	31241	9879
150	31278	31253	46900	15594	31248	31255	9900

*b) Lista następników*

LISTA NASTĘPNIKOW ALG. KAHNA							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	0	0	15628	0	0	3126	6251
40	15620	0	15621	15647	15623	12502	6251
50	15626	15624	15600	31247	15591	18738	6255
60	31249	31248	31256	31247	31271	31254	9
70	62503	62494	46904	46870	46870	53128	7651
80	78140	78123	78122	78148	62466	75000	6267
90	124960	109358	93771	109395	125011	112499	11680
100	156206	171829	187488	171864	171863	171850	9892
110	218700	218761	218732	234356	218736	221857	6250
120	296856	281254	328095	312475	359348	315606	26873
130	406224	390594	390601	374973	406219	393722	11692
140	499957	499963	499963	468744	484367	490599	12485
150	687483	624954	640551	640580	656170	649948	21205

LISTA NASTĘPNIKOW DFS							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	0	0	0	0	0	0	0
40	0	15657	0	0	0	3131	6263
50	0	0	0	0	0	0	0
60	0	15622	15598	0	0	6244	7647
70	15622	0	0	0	0	3124	6249
80	0	0	15629	0	15632	6252	7657
90	0	15625	15651	0	15625	9380	7659
100	15599	15625	15621	15625	15624	15619	10
110	15624	0	0	15594	0	6244	7647
120	15623	0	15654	15625	15623	12505	6253
130	15654	15622	31230	31222	31249	24995	7640
140	15622	15653	15626	15625	31256	18756	6250
150	31244	31273	31245	15599	15659	25004	7655

c) *Macierz grafu*

MACIERZ GRAFU ALG. KAHNA							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	31248	15649	15623	31217	15656	21879	7637
40	46838	46870	46903	31272	46845	43746	6237
50	93777	93712	78119	93774	93748	90626	6254
60	156210	140614	156272	156259	140613	149994	7659
70	218734	218733	250010	249952	234384	234363	13974
80	343757	328101	374971	359381	374945	356231	18211
90	499941	515557	499934	499937	499990	503072	6243
100	687453	656233	671797	703100	703051	684327	18216
110	937463	843688	874937	859312	859335	874947	32783
120	1093672	1171816	1218628	1203012	1124887	1162403	46969
130	1453022	1406150	1421772	1421739	1407049	1421946	16954
140	1749902	1812390	1765495	1703542	1832290	1772724	45784
150	2109216	2109188	2093623	2187338	2171685	2134210	37749

MACIERZ GRAFU DFS							
	1 pomiar	2 pomiar	3 pomiar	4 pomiar	5 pomiar	średnia	odchylenie
30	15649	18014	15624	31248	15649	19237	6076
40	62462	40976	31220	31247	46873	42556	11608
50	78119	73957	93776	78120	78119	80418	6871
60	187481	132950	140632	140614	140633	148462	19735
70	240533	190867	203111	234356	203136	214401	19439
80	328127	312508	281229	328074	312479	312483	17115
90	453059	437471	421844	421844	484308	443705	23370
100	593705	578055	609332	578083	593706	590576	11698
110	796841	765568	734321	765567	765534	765566	19771
120	1062393	984300	953052	953055	999896	990539	40249
130	1296778	1218660	1234310	1249907	1265531	1253037	26876
140	1687376	1515516	1562357	1546760	1546796	1571761	59778
150	1967038	1859239	1861078	1827993	1921767	1887423	50077

## 2. Zależność czasu obliczeń od liczby wierzchołków w grafie

a) Sortowanie topologiczne z wykorzystaniem Algorytmu Kahna

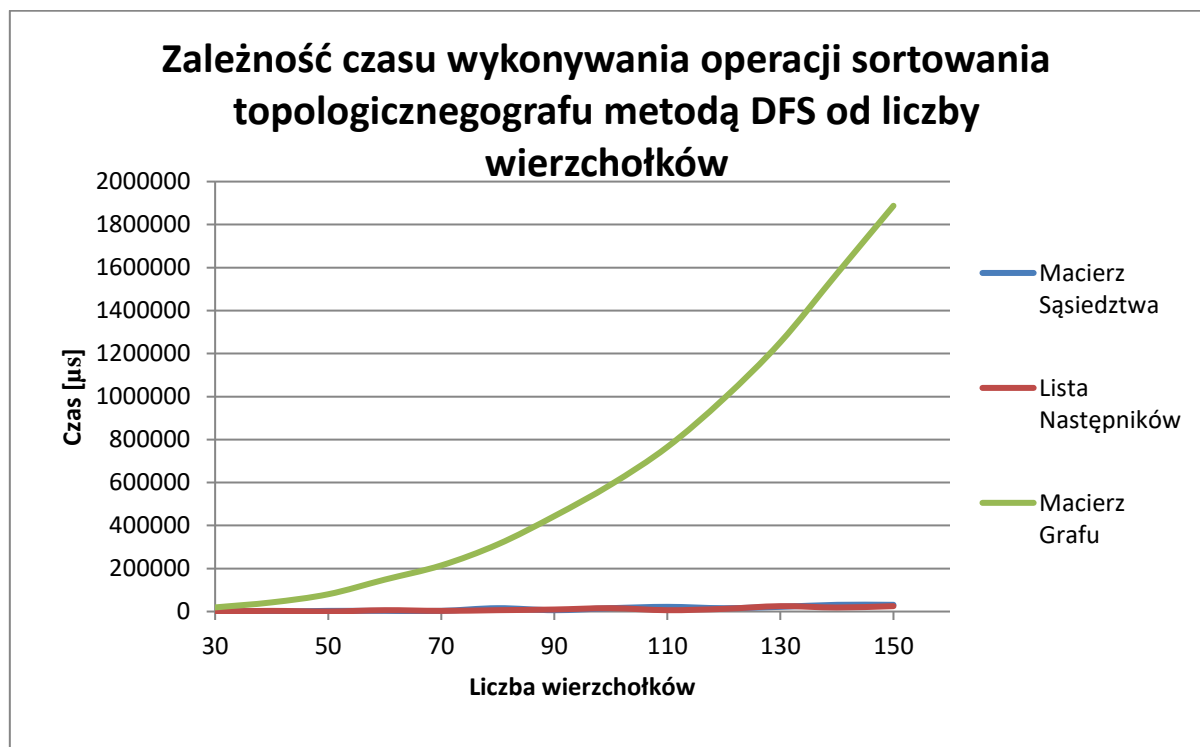
SORTOWANIE ALG. KAHNA			
	Macierz Sąsiedztwa	Lista Następników	Macierz Grafu
30	0	3126	21879
40	6254	12502	43746
50	3129	18738	90626
60	9368	31254	149994
70	15615	53128	234363
80	21867	75000	356231
90	31237	112499	503072
100	40613	171850	684327
110	46866	221857	874947
120	59376	315606	1162403
130	81249	393722	1421946
140	81256	490599	1772724
150	103121	649948	2134210



**OBSERWACJE:** Na wykresie wraz ze wzrostem liczby wierzchołków zarysowuje się różnica pomiędzy algorytmami. Przejście algorytmem Kahna najdłużej wykonuje się w przypadku macierzy grafu, następnie ze znacznie mniejszym wzrostem kolejno lista następników i macierz sąsiedztwa.

b) Sortowanie topologiczne z użyciem metody DFS

SORTOWANIE Z WYKORZYSTANIEM DFS			
	Macierz Sąsiedztwa	Lista Następników	Macierz Grafu
30	3124	0	19237
40	0	3131	42556
50	3123	0	80418
60	3120	6244	148462
70	3123	3124	214401
80	15625	6252	312483
90	6250	9380	443705
100	15629	15619	590576
110	21872	6244	765566
120	15628	12505	990539
130	21878	24995	1253037
140	31241	18756	1571761
150	31255	25004	1887423

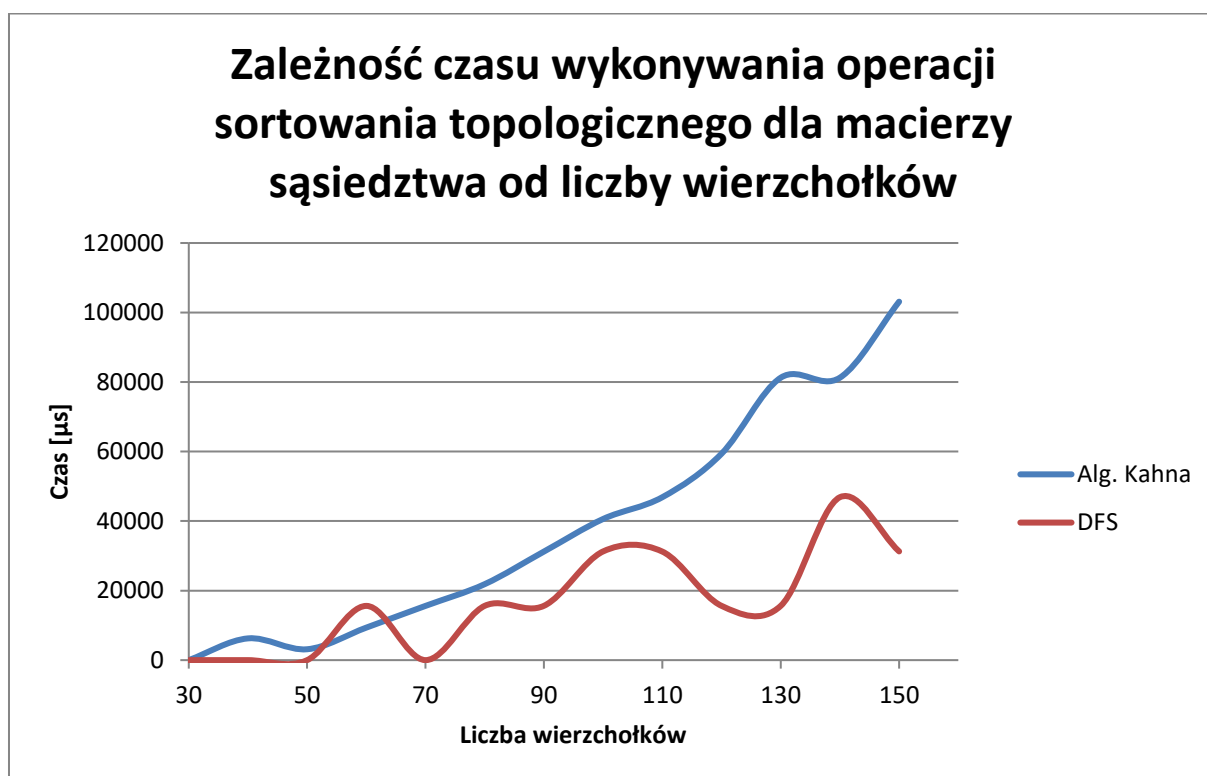


**OBSERWACJE:** Na wykresie znacznie zarysowuje się różnica pomiędzy reprezentacją maszynową w postaci macierzy grafu a listą następników i macierzą sąsiedztwa. Z naszych pomiarów wynika, że macierz grafu jest zdecydowanie mniej optymalna od pozostałych.

### 3. Zależność czasu sortowania topologicznego od liczby wierzchołków w grafie

a) Macierz sąsiedztwa

	Macierz Sąsiedztwa	
	ALG. KAHNA	DFS
30	0	0
40	6254	0
50	3129	0
60	9368	15601
70	15615	0
80	21867	15624
90	31237	15626
100	40613	31271
110	46866	31226
120	59376	15622
130	81249	15621
140	81256	46874
150	103121	31278



**OBSERWACJE:** Na wykresie wraz ze wzrostem liczby wierzchołków zarysowuje się różnica pomiędzy algorytmami. W przypadku macierzy sąsiedztwa bardziej optymalny okazuje się algorytm DFS w porównaniu do algorytmu Kahna.



b) Lista następników

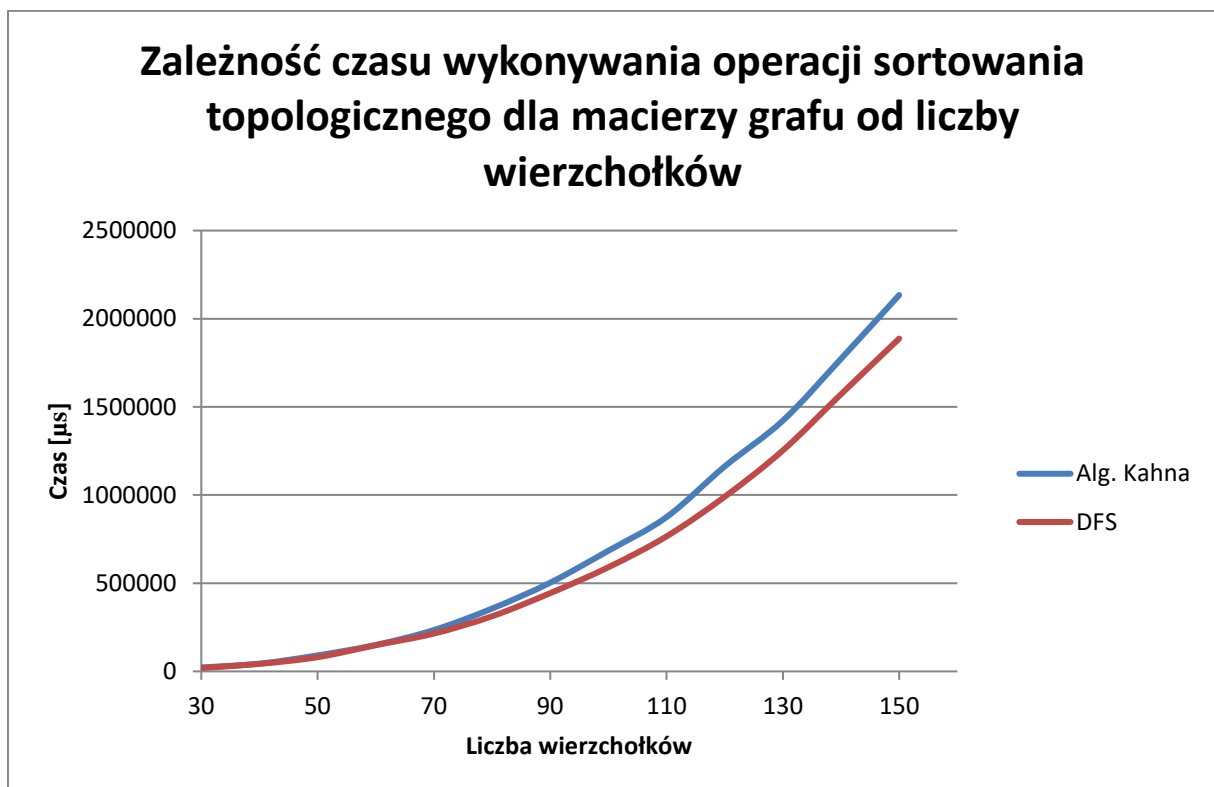
	Lista Następników	
	ALG. KAHNA	DFS
30	3126	0
40	12502	3131
50	18738	0
60	31254	6244
70	53128	3124
80	75000	6252
90	112499	9380
100	171850	15619
110	221857	6244
120	315606	12505
130	393722	24995
140	490599	18756
150	649948	25004



**OBSERWACJE:** Na wykresie można zauważyć, że algorytm DFS jest znacznie bardziej optymalny czasowo od algorytmu Kahna dla listy następników.

c) *Macierz Grafu*

	Macierz Grafu	
	ALG. KAHNA	DFS
30	21879	19237
40	43746	42556
50	90626	80418
60	149994	148462
70	234363	214401
80	356231	312483
90	503072	443705
100	684327	590576
110	874947	765566
120	1162403	990539
130	1421946	1253037
140	1772724	1571761
150	2134210	1887423



**OBSERWACJE:** Na wykresie różnica pomiędzy przejściem DFS i algorytmem Kahna jest niewielka. Algorytmy sortowania topologicznego dla macierzy grafu charakteryzują się bardzo podobną szybkością działania, jednak minimalnie szybciej działa metoda DFS.

## 4. Podsumowanie

Główną zaletą macierzy sąsiedztwa jest jej budowa, która pozwala łatwo odnieść się do konkretnej komórki w celu sprawdzenia, czy dany łuk istnieje. W przypadku przejścia metodą Kahna pozwala to w łatwy sposób odnaleźć wierzchołek o zerowym stopniu wejściowym poprzez sprawdzenie, czy w danym wierszu znajduje się wartość „-1” i następnym ewentualnym wyzerowaniu wartości danego wiersza i kolumny. Sytuacja prezentuje się podobnie przy przejściu DFS, przy przejściu rekurencyjnym sprawdzamy warunek, czy w danym wierszu znajduje się wartość „1”, więc czy należy odwiedzić kolejny wierzchołek. Ponadto, tworzenie macierzy sąsiedztwa jest łatwe w implementacji, co również możemy uznać jako zaletę tej reprezentacji grafu.

Aby zdefiniować macierz sąsiedztwa potrzebna jest macierz kwadratowa, do której wpisywane zostają poszczególne łuki (złożoność pamięciowa wynosi  $O(V^2)$ ). Przekłada się to na potrzebę zarezerwowania dużej przestrzeni w pamięci, co w przypadku grafów z dużą liczbą wierzchołków i relatywnie małą liczbą krawędzi może okazać się wadą.

W przypadku listy następników potrzebujemy listę każdego następnika danego wierzchołka. Zatem, jego główną zaletą okazuje się optymalizacja potrzebnej przestrzeni w pamięci, zwłaszcza przy małej liczbie krawędzi. Ponadto, tworzenie listy następników jest łatwe w implementacji.

Zdecydowaną wadą listy następników okazuje się w przypadku przejścia algorytmem Kahna potrzeba przeiterowania się po każdym wierzchołku w celu sprawdzenia, czy dany wierzchołek nie jest następnikiem innego. Dodatkowo każdy z tych wierzchołków może posiadać wiele następników i każdy z nich musimy porównać z danym wierzchołkiem do ewentualnego usunięcia. Sytuacja zmienia się w przypadku przejścia metodą DFS. Algorytm ten w przeciwieństwie do algorytmu Kahna rozpoczyna przejście rekurencyjne od pierwszego wierzchołka, który nie został dodany na stos, a ewentualne następne przejście rozpoczynamy od pierwszego następnika danego wierzchołka, do którego wystarczy się odwołać. Stąd na wykresie zarysowuje się wyraźna różnica pomiędzy algorytmami Kahna i DSF dla listy następników.

Główną zaletą macierzy grafu okazuje się ilość informacji które przechowuje. Pierwszy następnik, poprzednik i pierwszy nieincydentny wierzchołek danego wierzchołka to dużo informacji potrzebnych do iterowania się po grafie.

Wadą algorytmu, podobnie jak w przypadku macierzy sąsiedztwa okazuje się przestrzeń pamięciowa potrzebna do zdefiniowania macierzy. W tym przypadku potrzebujemy dodatkowo trzy kolumny.

Podsumowując, najbardziej optymalną reprezentacją grafu w postaci maszynowej okazuje się macierz sąsiedztwa. Jej łatwość implementacji i szybkość przetwarzania danych w porównaniu do listy następników i macierzy grafu jest znaczna.