

Politechnika Wrocławska
Wydział Informatyki i Telekomunikacji



Sieci złożone

Sprawozdanie z laboratorium

Autor
Maciej Majewski
nr albumu: 254295
kierunek: Inżynieria systemów
28 styczeń 2022

1. Wstęp

W poniżej pracy podjąłem się analizy danych lotniczych w Stanach Zjednoczonych z pierwszej połowy 2015 roku. Zająłem się analizą podstawowych właściwości stworzonej sieci złożonej oraz zbadałem wpływ odporność sieci na usuwania połączeń dobranych celowo do skutecznego rozspójnienia sieci.

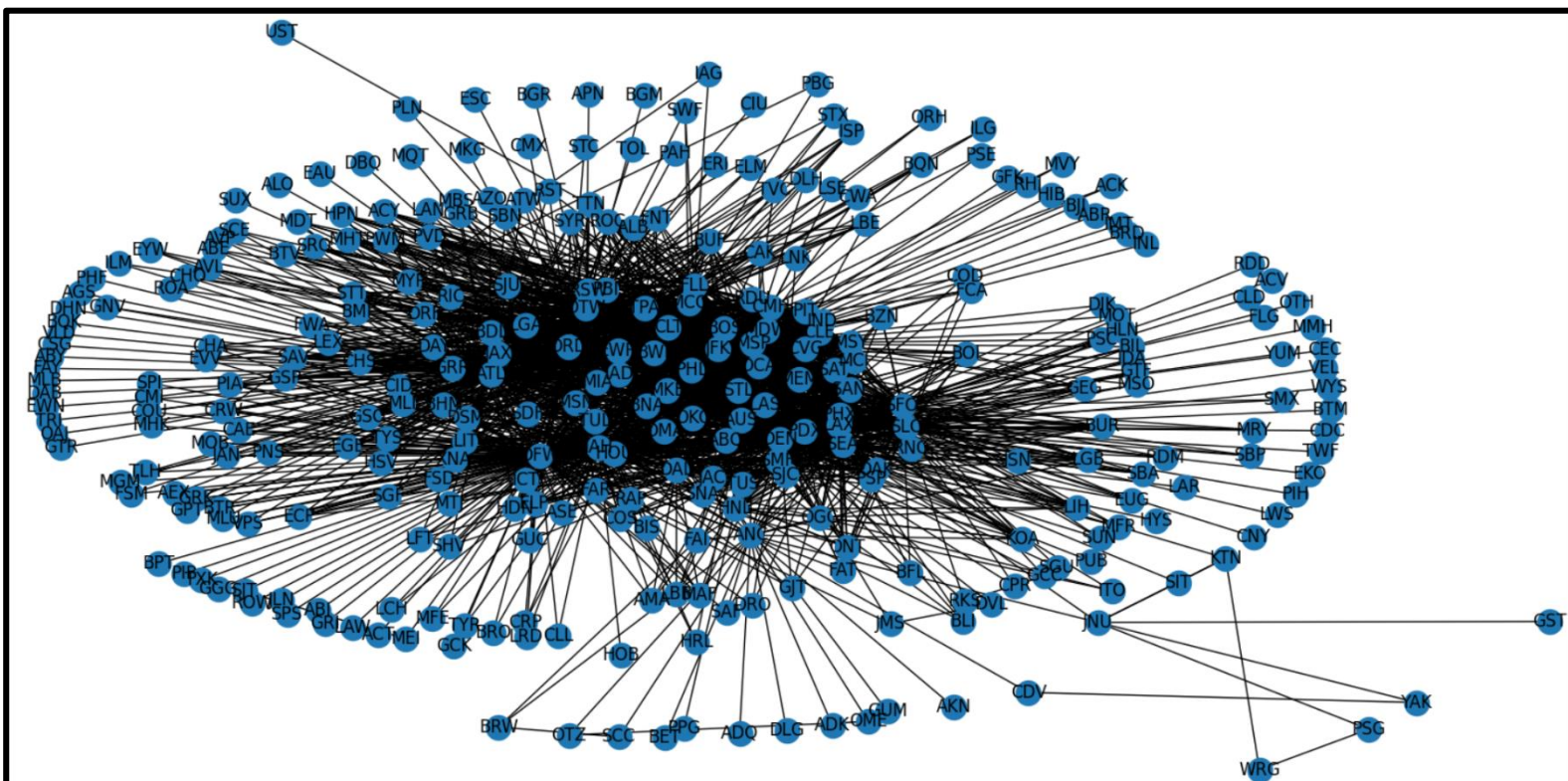
2. Opis rozwiązania

I. Dane

Dane na temat lotów w Stanach Zjednoczonych z pierwszej połowy 2015 roku zostały udostępnione przez Departament Transportu na stronie <https://www.kaggle.com/usdot/flight-delays>. Pobrana baza danych zawiera daną z całego roku, natomiast w związku z bardzo dużą liczbą rekordów, ograniczyłem się do pierwszej połowy roku oraz tylko do interesujących mnie cech danego połączenia. Do obsługi tych danych wykorzystałem bibliotekę *Pandas*.

II. Wizualizacja sieci złożonej

Załadowane i wstępnie przetworzone dane zostały zwizualizowane w formie grafu za pomocą bibliotek *Matplotlib* oraz *Networkx*. Graf (Rysunek 1) pokazuje realizowane połączenia pomiędzy poszczególnymi lotniskami na terenie Stanów Zjednoczonych. W związku z długimi nazwami lotnisk, węzły opisane zostały typowymi skrótami kodu IATA. Sama wizualizacja nie jest czytelna w związku z dużą ilością węzłów, natomiast daje nam wstępny pogląd jak wygląda rozkład węzłów w sieci złożonej. Widzimy dobrze centrum posiadające większy stopień oraz obrzeże mające zaledwie kilka połączeń.



Rysunek 1 - Graf połączeń lotniczych

3. Rezultaty obliczeń

I) Podstawowe wielkości sieci

- 1) Ilość węzłów w sieci = **320**, czyli lotnisk, które zostały uwzględnione w bazie danych i występują pomiędzy nimi połączenia.
- 2) Ilość połączeń w sieci = **2156**, czyli liczba połączeń pomiędzy dwoma lotniskami, które zostały uwzględnione w bazie danych.
- 3) Gęstość sieci \approx **0.0422**, czyli wartość pokazująca jak duży jest stosunek aktualnych połączeń w sieci do wszystkich możliwych połączeń. Wartość ta mieści się w zakresie $[0,1]$, gdzie 0 opisuje graf bez połączeń, natomiast 1 – graf pełny.
- 4) Średnica = **5**, czyli długość najdłuższej z najkrótszych ścieżek w grafie. Wartość ta oznacza, że dla dowolnie wybranych dwóch wierzchołków, ilość połączeń lotniczych jakie należy pokonać wynosi maksymalnie (w tym wypadku) 5.
- 5) Najważniejsze połączenia:

Nr.	Połączenie z	Połączenie do	Współczynnik centralności
1	Seattle (SEA)	Juneau (JNU)	0.0147
2	Anchorage (ANC)	Atlanta (ATL)	0.0125
3	Seattle(SEA)	Ketchikan (KTN)	0.0120
4	Anchorage (ANC)	Chicago (ORD)	0.0113
5	Anchorage (ANC)	Dallas-Fort Worth (DFW)	0.0107
6	Atlanta (ATL)	Salt Lake City (SLC)	0.0102
7	Dallas-Fort Worth (DFW)	Atlanta (ATL)	0.0098
8	Dallas-Fort Worth (DFW)	Salt Lake City (SLC)	0.0095
9	Dallas-Fort Worth (DFW)	Chicago (ORD)	0.0094
10	Dallas-Fort Worth (DFW)	Salt Lake City (SLC)	0.0092

Tabela 1 - Najważniejsze połączenia lotnicze

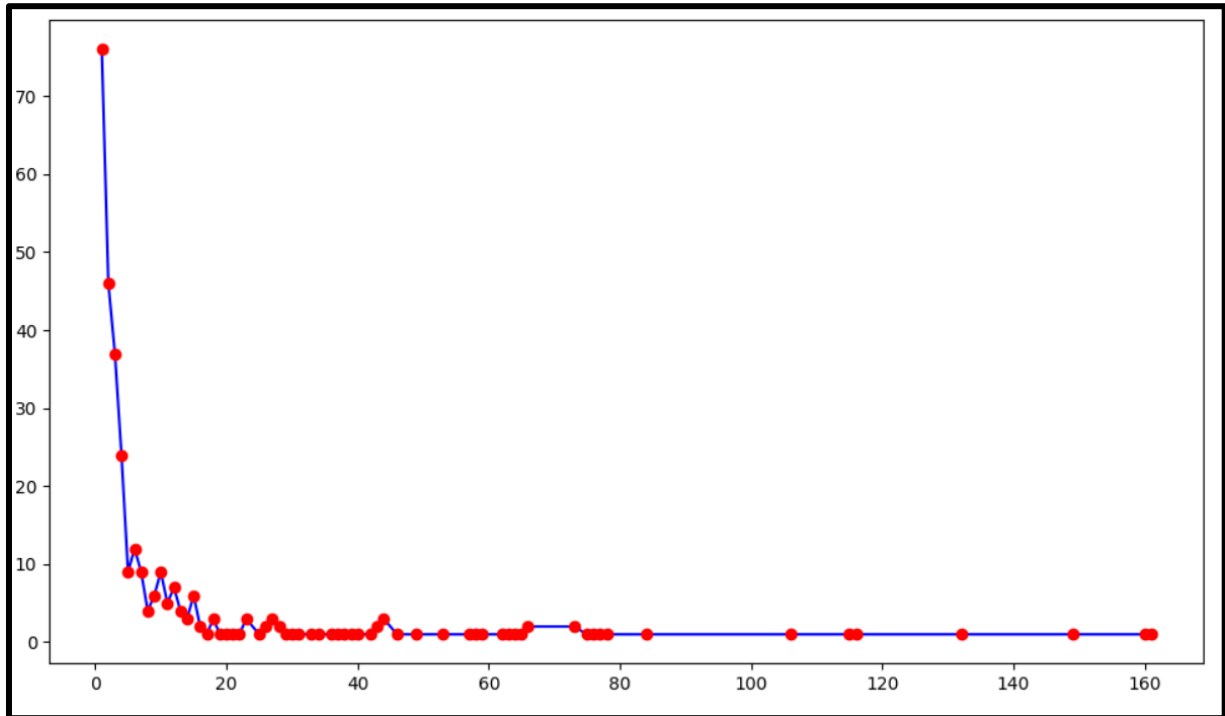
- 6) Najważniejsze lotniska:

Nr.	Lotnisko	Współczynnik centralności	Ilość połączeń
1	Atlanta (ATL)	0.5047	161
2	Chicago (ORD)	0.5016	160
3	Dallas-Fort Worth (DFW)	0.4671	149
4	Denver (DEN)	0.4138	132
5	Minneapolis (MSP)	0.3636	116
6	Houston (IAH)	0.3605	115
7	Detroit (DTW)	0.3323	106
8	Salt Lake City (SLC)	0.2633	84
9	Newark (EWR)	0.2445	78
10	Los Angeles (LAX)	0.2414	77

Tabela 2 - Najważniejsze lotniska

II) Rozkład węzłów

Rozkład węzłów w sieci lotniczej został zwizualizowany na Rysunku 2. Po przedstawieniu stopni wszystkich węzłów można zauważyć, że ich rozkład może być opisany rozkładem potęgowym.



Rysunek 2- Rozkład węzłów sieci

III) Analiza połączeń

Opis analizy

Podjąłem się przetestowania sieci połączeń lotniczych pod względem jej odporności na usuwanie z niej połączeń. Połączenia z sieci były usuwane pojedynczo tak długo aż był spełniony warunek spójności sieci. Pozwoliło to na sprawdzenie ile połączeń musiałoby zostać zawieszonych, aby nie było możliwości dostania się z dowolnego lotniska na inne.

Do przetestowania odporności sieci wykorzystałem metodę podziałową, to znaczy usuwałem połączenia, które posiadały największy współczynnik centralności w danej chwili. Takie podejście pozwala skutecznie rozspójnić sieć, gdyż usuwa się najważniejsze połączenia.

a) Usuwanie najważniejszych połączeń

W przypadku zastosowania podejścia zakładającego usuwanie połączeń o największym współczynnik centralności, sieć przestaje być spójna dość szybko. Należy usunąć zalewie 4 takie połączenia. Warto zaznaczyć, że wszystkie te połączenia, łączą przedzielony Kanadą stan Alaski z resztą kraju. Położenie tego stanu, oraz mała liczba połączeń pomiędzy lotniskami w tym stanie z resztą sprawia, że bardzo łatwo go odłączyć.

Nr.	Połączenie z	Połączenie do	Współczynnik centralności	Średnica sieci
1	Seattle (SEA)	Juneau (JNU)	0.0147	5
2	Anchorage (ANC)	Juneau (JNU)	0.0210	6
3	Seattle (SEA)	Ketchikan (KTN)	0.0224	6
4	Seattle (SEA)	Sitka (SIT)	0.0365	8

Tabela 3 - Usuwane połączenia lotnicze

b) Usuwanie najważniejszych połączeń – wersja 2

W związku z położeniem stanu Alaski i jego lotnisk postanowiłem przeprowadzić test ponownie, jednakże usuwając wcześniej wszystkie połączenia lotnicze związane z lotniskami na Alasce. Dzięki temu w bazie posiadamy tylko lotniska znajdujące się w pozostałych stanach.

Podejście to przyniosło oczekiwany efekt, ponieważ w tym przypadku do rozspójnienia sieci należało usunąć aż 30 połączeń, które zostały przedstawione na Rysunku 3. W tym przypadku połączenia obejmują różne lotniska na terenie kraju.

0 ['SLC', 'ATL']	1 ['DFW', 'SLC']
2 ['ORD', 'SLC']	3 ['DEN', 'ATL']
4 ['DFW', 'ATL']	5 ['ORD', 'DFW']
6 ['DFW', 'MSP']	7 ['DFW', 'DEN']
8 ['ORD', 'DEN']	9 ['ATL', 'MSP']
10 ['ORD', 'ATL']	11 ['DFW', 'DTW']
12 ['SLC', 'IAH']	13 ['SLC', 'DTW']
14 ['ORD', 'MSP']	15 ['DEN', 'DTW']
16 ['ORD', 'IAH']	17 ['MSP', 'IAH']
18 ['DEN', 'IAH']	19 ['SFO', 'ATL']
20 ['SFO', 'DFW']	21 ['SFO', 'ORD']
22 ['SLC', 'MSP']	23 ['DEN', 'MSP']
24 ['DTW', 'MSP']	25 ['ATL', 'DTW']
26 ['DTW', 'IAH']	27 ['ATL', 'IAH']
28 ['SFO', 'DTW']	29 ['ATL', 'ABY']

Rysunek 3 - Usunięte połączenia

4. Wnioski

Po wykonaniu zadania, którym była analiza sieci, mogę stwierdzić, że sieć połączeń lotniczych w Stanach Zjednoczonych jest siecią złożoną. Jest siecią o złożonej strukturze, dużej liczbie węzłów i posiadającą węzły spełniające funkcję hubów.

Dodatkowo po przeprowadzeniu testów odporności sieci na usunięcia połączeń mogę stwierdzić, że sieć połączeń można łatwo rozspójnić w związku z ułożeniem kraju jakim są Stany Zjednoczone. Usunięcie z testów stanu Alaska uodparnia sieć złożoną na usuwanie połączeń.