



VYTAUTO DIDŽIOJO UNIVERSITETAS

INFORMATIKOS FAKULTETAS

TAIKOMOSIOS INFORMATIKOS KATEDRA

Socialinių medijų analitika (INF6018)

Europos oro uostų tinklo analizė: geografijos įtaka skrydžių srautams

Atliko: Edvinas Merkevičius MIF240015

Laboratorinių darbų dėstytojas (- ai): Judita Kasperiūnienė

Kaunas, 2025

TURINYS

1.	ĮVADAS	1
2.	METODAI IR DUOMENYS	1
3.	REZULTATAI.....	3
4.	IŠVADOS IR INTERPRETACIJA	6
5.	LITERATŪROS ŠALTINIAI	7

1. ĮVADAS

Projekto aprašymas: Šiame darbe atliekama 71 pagrindinio Europos oro uosto tinklo analizė. Projektu siekiama ištirti Europos aviacijos struktūrą, identifikuoti svarbiausius transporto mazgus („Hubs“) ir nustatyti, kokie veiksniai lemia skrydžių srautų formavimąsi.

Tyrimo klausimas: Pagrindinis darbo klausimas: Ar Europos skrydžių tinklas formuojasi grynai geografiniu principu (regioniniai klasteriai), ar struktūrą lemia kiti faktoriai? Taip pat keliama hipotezė, kad nepaisant globalizacijos, Europos aviacija išlieka stipriai segmentuota į regionines bendruomenes (pvz., Šiaurės, Pietų, Vakarų blokai), o mažesni oro uostai (pvz., Vilnius) veikia tik kaip periferiniai taškai.

2. METODAI IR DUOMENYS

Duomenų šaltinis: Analizei naudojami realūs aviacijos duomenys iš atvirojo kodo projekto „OpenFlights“ (šaltinis: *openflights.org*). Duomenų rinkmena suformuota programiškai, naudojant Python biblioteką *pandas* tiesioginiam duomenų nuskaitymui.

Duomenų apdorojimas: Pradinis duomenų rinkinys apėmė virš 10 000 pasaulio oro uostų. Siekiant suformuoti analizei tinkamą tinklą (50–100 mazgų), atlikti šie filtravimo žingsniai:

1. Duomenų nuskaitymas.

```
airports_url = "https://raw.githubusercontent.com/.../airports.dat"
routes_url = "https://raw.githubusercontent.com/.../routes.dat"
df_airports = pd.read_csv(airports_url, header=None)
df_routes = pd.read_csv(routes_url, header=None)
```

2. Atrinkti tik Europos šalių oro uostai.

```
europe_countries = ["Lithuania", "Germany", "United Kingdom", "France",
                    "Spain", ...]
europe_airports =
df_airports[df_airports['Country'].isin(europe_countries)]
```

3. Pagal skrydžių intensyvumą atrinkta 70 didžiausių mazgų (TOP 70).

```
active_airports = df_routes['SourceAirport'].value_counts()
top_70_ids = active_airports.head(70).index.tolist()
```

4. Įtrauktas Vilniaus oro uostas (VNO) palyginimui.

```
if 'VNO' not in top_70_ids: top_70_ids.append('VNO')
```

5. Suformuotos briaunos (Edges) tik tarp šių atrinktų oro uostų.

```
final_routes = df_routes[
    (df_routes['SourceAirport'].isin(top_70_ids)) &
    (df_routes['DestAirport'].isin(top_70_ids))
]
```

6. Jei kelios avialinijos skrenda tuo pačiu maršrutu, mums tai viena briauna.

```
final_edges = final_routes[['SourceAirport',
    'DestAirport']].drop_duplicates()
```

7. Galutinis išsaugojimas.

```
final_nodes = europe_airports[europe_airports['IATA'].isin(top_70_ids)]
final_nodes.to_csv("openflights_nodes.csv", index=False)
final_edges.to_csv("openflights_edges.csv", index=False)
```

Tinklo elementai:

- **Mazgai (Nodes):** 71 Oro uostas.
- **Briaunos (Edges):** 2895 tiesioginiai skrydžiai.
- **Atributai:** Geografinės koordinatės (Platuma, Ilguma) ir Laiko juosta.

Naudoti algoritmai (Kodo fragmentai): Analizė atlikta „Google Colab“ aplinkoje naudojant Python.

1. Tinklo kūrimas:

```
# Duomenų nuskaitymas ir tinklo formavimas
df_nodes = pd.read_csv('openflights_nodes.csv')
df_edges = pd.read_csv('openflights_edges.csv')
G = nx.from_pandas_edgelist(df_edges, source='Source', target='Target')
```

2. Svarbos nustatymas (Centrality):

```
# Skaiciuojame populiarumą ir tarpininkavimą
degree = nx.degree_centrality(G)
betweenness = nx.betweenness_centrality(G)
```

3. **Bendruomenių aptikimas (Louvain):** Naudotas Louvain algoritmas, kuris optimizuoja tinklo moduliariškumą, ieškodamas tankių grupių.

```
import community.community_louvain as community_louvain
partition = community_louvain.best_partition(G)
```

4. **Klasterizavimas (K-Means):** Oro uostai sugrupuoti pagal geografines koordinates, naudojant *KMeans* algoritmą (3 klasteriai).

```
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler

# Pasirenkame duomenis (koordinates ir laiko juosta)
```

```
features = df_nodes[['Latitude', 'Longitude', 'Timezone']].fillna(0)

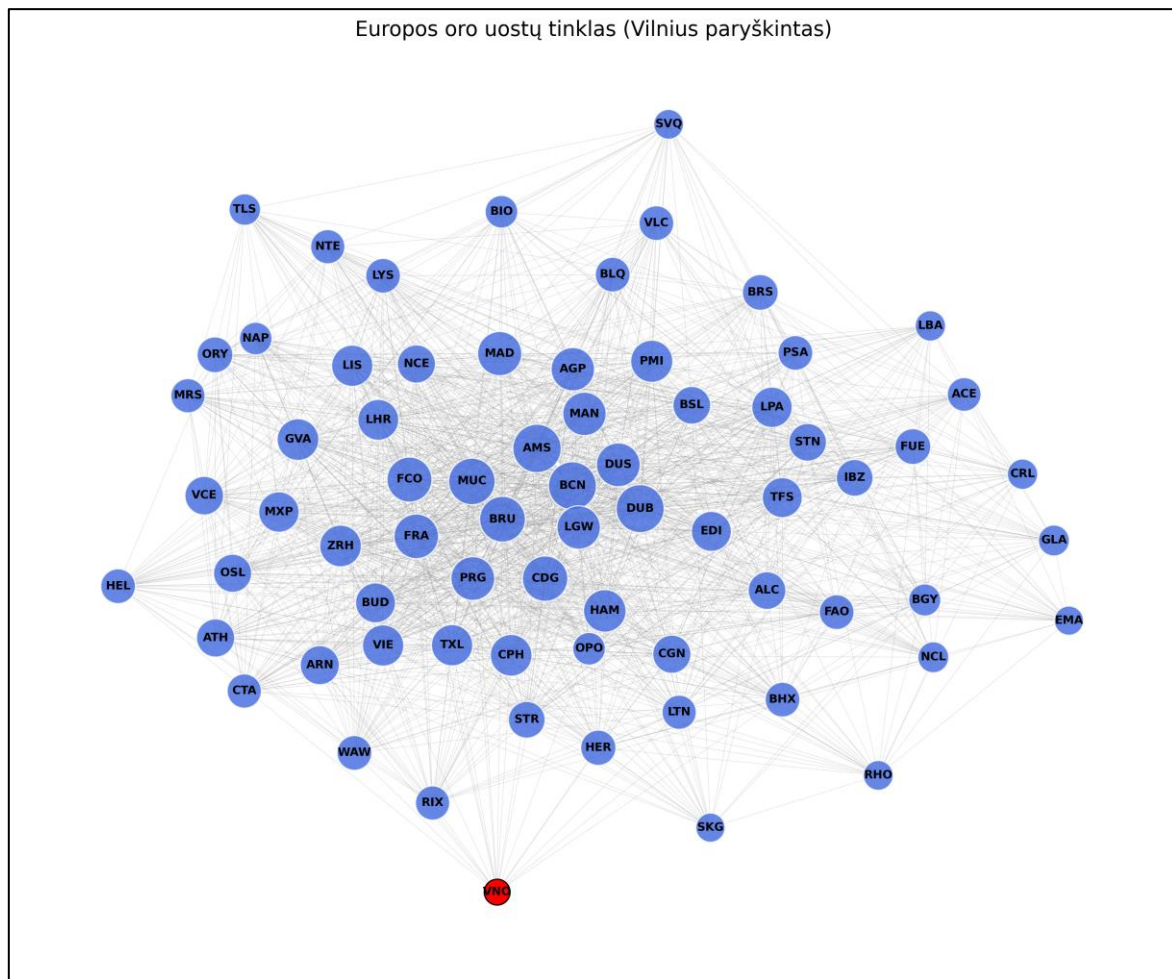
# Standartizuojuame duomenis (suvienodiname mastelį)
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features)

# Sugrupuojame į 3 geografinius regionus
kmeans = KMeans(n_clusters=3, random_state=42)
df_nodes['Cluster_Geo'] = kmeans.fit_predict(features_scaled)
```

3. REZULTATAI

Tinklo parametrų analizė: Atlikus skaičiavimus, nustatyti svarbiausi Europos aviacijos centrai.

- **Populiariausi mazgai (Degree Centrality):** Didžiausią tiesioginių skrydžių skaičių tinkle turi **Amsterdamas (AMS)** (rodiklis: 0.94). Tai rodo, kad šis oro uostas yra pagrindinis Europos „Hubas“. Jam iš paskos seka **Dublinas (DUB)** ir **Barselona (BCN)** (abu po 0.93).
- **Tarpininkai (Betweenness Centrality):** Pagal tarpininkavimo rodiklį taip pat pirmauja **Dublinas (DUB)** (0.0188). Tai rodo jo strateginę svarbą jungiant skirtingus regionus (tikėtina, dėl pigių skrydžių bendrovių tinklo).
- **Vilniaus (VNO) pozicija:** Vilniaus oro uostas šiame tinkle užima **71-ąją vietą** (iš 71). Jo tiesioginių skrydžių rodiklis yra žemas (0.30), o tarpininkavimo rodiklis artimas nuliui. Tai patvirtina, kad Vilnius yra periferinis mazgas, priklausomas nuo didžiųjų centrų.

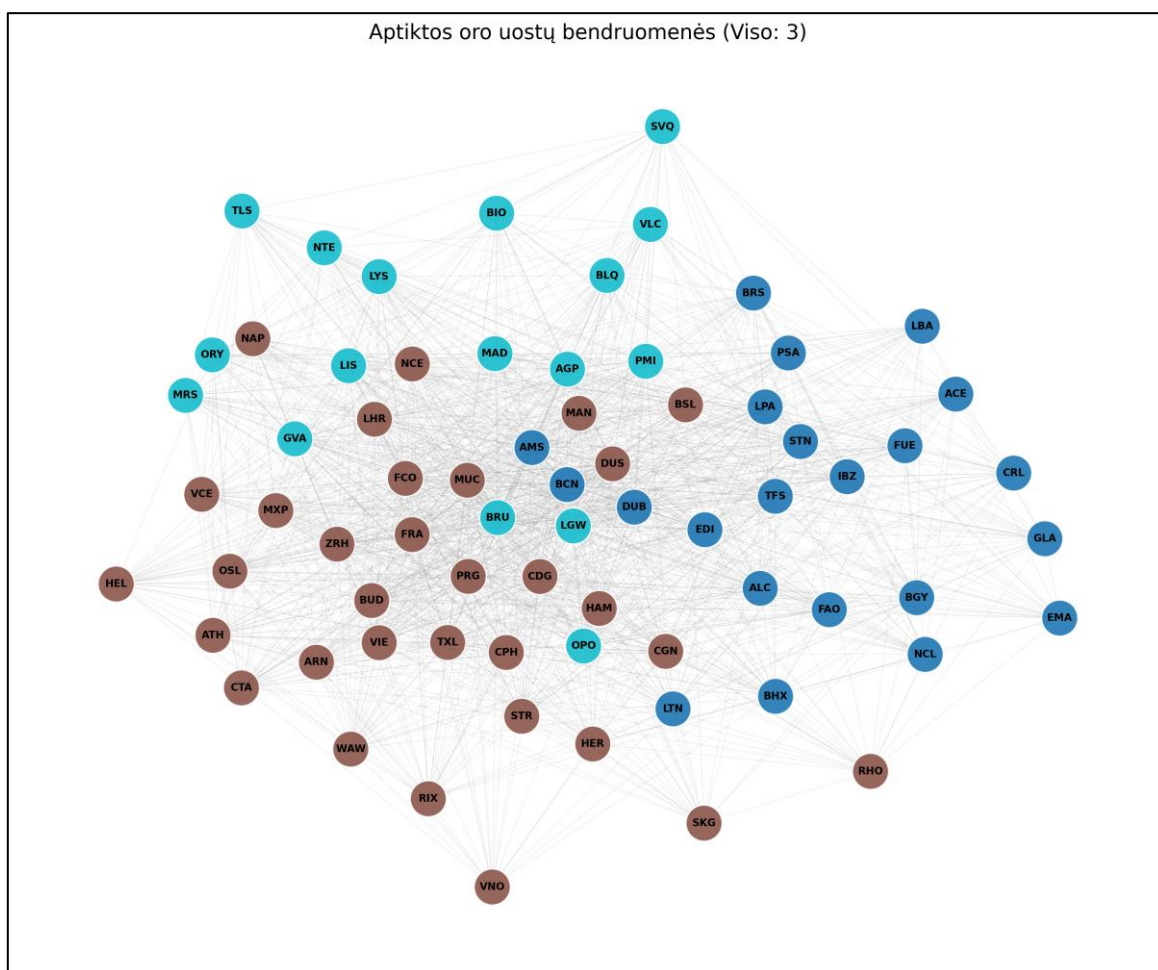


pav. 1 Europos oro uostų tinklas pagal svarbą. Raudonas taškas – Vilnius

Bendruomenių aptikimas: Naudojant Louvain algoritimą, Europos oro uostų tinkle buvo identifikuotos 3 pagrindinės bendruomenės. Išanalizavus jų sudėtį, matyti, kad tinklas skyla ne tik geografiniu, bet ir funkcinio (skrydžių tipo) principu:

- **1 Bendruomenė (Pietvakarių Europa):** Šioje grupėje (šviesiai mėlyna) dominuoja **Ispanijos, Portugalijos ir Prancūzijos** oro uostai (MAD, LIS, ORY, MRS). Tai rodo glaudų regioninį susisiekimą tarp Iberijos pusiasalio ir Prancūzijos. Čia taip pat patenka Briuselis (BRU) ir Ženeva (GVA), kurie veikia kaip šio regiono vartai į šiaurę.
- **2 Bendruomenė (Pigių skrydžių ir atostogų kryptys):** Ši grupė (mėlyna) yra labai specifinė. Joje dominuoja **Jungtinės Karalystės ir Airijos** oro uostai (STN, LTN, DUB, EDI) bei populiariausios **atostogų kryptys** (Alikantė, Tenerifė, Faro, Ibiza). **Interpretacija:** Šią bendruomenę formuoja pigių skrydžių bendrovių (pvz., „Ryanair“, „EasyJet“) maršrutai, kurie jungia Britų salas su pietų kurortais. Čia taip pat patenka Amsterdamas (AMS), veikiantis kaip pagrindinis srautų skirstytuvas.

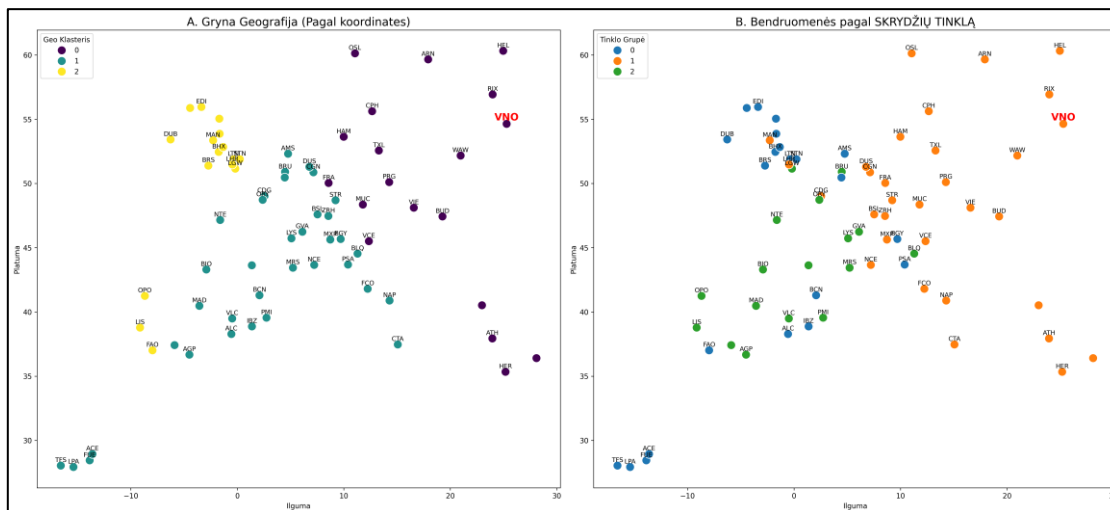
- 3 Bendruomenė (Centrinė ašis ir Rytų/Šiaurės Europa):** Tai didžiausia ir „politinė“ grupė (ruda). Jos branduolį sudaro **Vokietijos, Italijos ir Centrinės Europos** didieji centrai („Hubai“): Frankfurtas (FRA), Miunchenas (MUC), Ciurichas (ZRH), Viena (VIE). Svarbu tai, kad ši grupė surenka srautus iš **Skandinavijos** (OSL, ARN, HEL) ir **Rytų Europos/Baltijos šalių** (WAW, RIX, VNO). **Lietuvos pozicija:** Vilniaus oro uostas (VNO) priklauso šiai bendruomenei, kas patvirtina jo priklausomybę nuo tradicinių vežėjų (pvz., „Lufthansa“, „airBaltic“, „SAS“) maršrutų per Frankfurtą, Kopenhagą ar Rygą.



pav. 2 Oro uostų bendruomenės pagal skrydžių srautus

4. IŠVADOS IR INTERPRETACIJA

Lyginamoji analizė (Tinklas vs. Geografija): Siekiant atsakyti į tyrimo klausimą, buvo atliktas vizualus palyginimas tarp grynai geografinio grupavimo (K-Means pagal koordinatas) ir realių skrydžių bendruomenių.



pav. 3 Kairėje – Geografiniai klasteriai, Dešinėje – Skrydžių tinklo bendruomenės

Pagrindinės išvados:

1. **Geografinis ir funkcinis determinizmas:** Lyginant A ir B grafikus (Pav. 3), matoma stipri koreliacija, tačiau su svarbiu niansu. Nors Pietų ir Vidurio Europos bendruomenės formuojasi **geografiniu principu** (šalys jungiasi su kaimynėmis), Jungtinės Karalystės ir Airijos klasteris išsiskiria **funkciniu ryšiu**. Jis jungia ne tik kaimynines šalis, bet ir nutolusius „atostogų“ taškus (Ispanijos salas, Portugaliją), kas rodo stiprią pigių skrydžių bendrovių („Ryanair“, „EasyJet“) įtaką tinklo struktūrai.
2. **Tinklo struktūra:** Tinklas nėra centralizuotas aplink vieną tašką, bet turi kelis dominuojančius centrus – **Amsterdamą, Dubliną, Frankfurtą**. Šie „Hubai“ atlieka skirtingus vaidmenis: Frankfurtas ir Miunchenas veikia kaip tradiciniai tranzito centrai, surenkantys srautus iš regionų (pvz., Rytų Europos), o Dublynas ir Barcelona dominuoja tiesioginių ("taškas-taškas") skrydžių segmente.
3. **Atsakymas į klausimą (Lietuvos pozicija):** Tyrimas parodė, kad geografinė padėtis yra pagrindinis veiksnys, lemiantis oro uosto vietą tinkle. Vilnius, būdamas geografinėje periferijoje, natūraliai atsiduria ir tinklo pakraštyje. Jis neturi aukšto tarpininkavimo rodiklio ir veikia kaip „spoke“ (galutinis taškas), kurio susisiekimas priklauso nuo jungčių su didžiaisiais Centrinės Europos (Frankfurto, Rygos, Varšuvos) mazgais.

5. LITERATŪROS ŠALTINIAI

1. GitHub Projektas: <https://github.com/Galahexolion/SMA-skrydziai>
2. OpenFlights Airports Data:
<https://raw.githubusercontent.com/jpatokal/openflights/master/data/airports.dat>
3. OpenFlights Routes Data:
<https://raw.githubusercontent.com/jpatokal/openflights/master/data/routes.dat>
4. NetworkX Documentation: <https://networkx.org/>
5. Blondel, V. D., et al. (2008). "Fast unfolding of communities in large networks".
6. Scikit-learn (K-Means Clustering): <https://scikit-learn.org/>
7. Sun, X., & Wandelt, S. (2021). "The structure of the global air transport network: A review from the perspective of complex networks".
8. Lordan, O., Sallan, J. M., & Simo, P. (2016). "Study of the topology and robustness of airline route networks".
9. Dai, L., Derudder, B., & Liu, X. (2018). "The evolving structure of the World Airline Network: Global communities and regional cores".
10. Zanin, M., & Lillo, F. (2013). "Modelling the air transport with complex networks: A short review".