Final Project Report

Flights Route Analysis :

Link Prediction and Optimization

Group 26

Group Member

Group 26

M11015018	M11015080	M11015064
陳彥家	湯傑堯	何昆霖

Outlines

- Dataset

- Method

- Results

- Conclusion

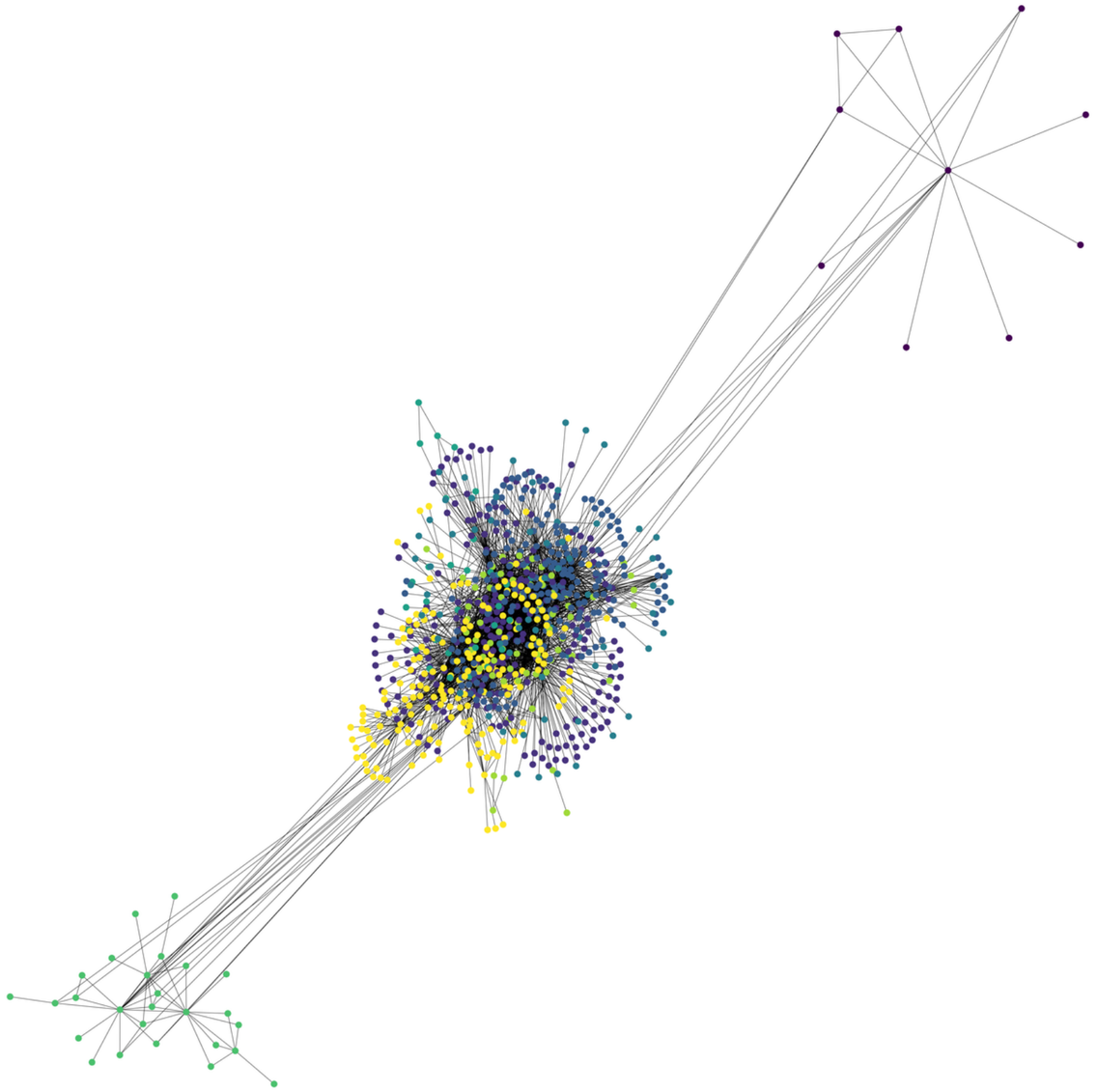
Dataset

- OpenFlight https://openflights.org/data.html

Contents

Routes	Airports	Airlines
59036	3209	531

Dataset



Dataset - OpenFlight

- Routes.dat Contains the routes of airlines

Airline, AID, Source, SID, Desitination, DID, codeshare, stops, equipments

	Airline	AID	Source	SID	Desitination	DID	codeshare	stops	equipments
0	2B	410	AER	2965	KZN	2990	NaN	0	CR2
1	2B	410	ASF	2966	KZN	2990	NaN	0	CR2
2	2B	410	ASF	2966	MRV	2962	NaN	0	CR2
3	2B	410	CEK	2968	KZN	2990	NaN	0	CR2
4	2B	410	CEK	2968	OVB	4078	NaN	0	CR2

Dataset - OpenFlight

- Airlines.dat Contains the information of airlines

Airline, Airline ID, Source Airport, Source Airport ID, Dest Airport, Dest Airport ID, Codeshare, Stops, equipment

	AirlineID	Name	Alias	IATA	ICAO	Callsign	Country	Active
0	-1	Unknown	\N	-	NaN	\N	\N	Y
1	1	Private flight	\N	-	NaN	NaN	NaN	Y
3	3	1Time Airline	\N	1T	RNX	NEXTIME	South Africa	Y
10	10	40-Mile Air	\N	Q5	MLA	MILE-AIR	United States	Y
13	13	Ansett Australia	\N	AN	AAA	ANSETT	Australia	Y

Method

- Airline Metric
- Cost function
- Airline analyze

Method

Airline Metric

- number of edges in the airline $E(A)$

- change in effective diameter $\Delta g = g(G \cup A) - g(G)$

- change in the clustering coefficient $\Delta Cl = Cl(G \cup A) - Cl(A)$

- modularity $Q = Q(y)$

- change in the average closeness centrality $\Delta CC = CC(G \cup A) - CC(A)$

- change in the nodes $\Delta V = V(G \cup A) - V(G)$

Method

Cost Function

- Use CVX to find c_i

$$c(G, A) = \sum_{i=1}^{5} c_i X_i$$

Method

Algorithm

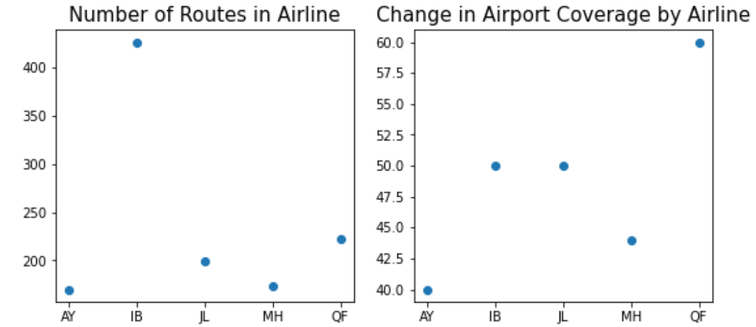
Algorithm 1 Feature Weighting

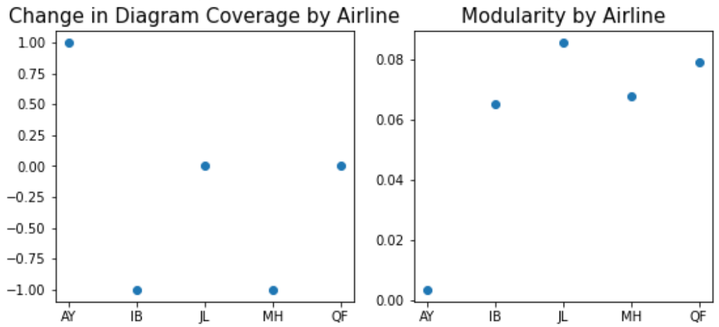
Input: $A = \{A_1, \ldots, A_N\}$ a set of airline graph
G an alliance graph
for $i = 1$ to N **do**
Remove A_i from G
end for
$F = N \times 6$ zero matrix
for $i = 1$ to N **do**
for $j = 1$ to 6 **do**
F_{ij} = value of the j^{th} metric with current G and A_i
end for
$G = G \cup A_i$
end for
for $j = 1$ to 6 **do**
Normalize the j^{th} column of F by its column standard deviation σ_j
end for
Solve in CVX:
Let $c = (c_1, \ldots, c_6)$
Solve: $F_C > 0$
Subject to: $\sum_{i=1}^{6} c_i = 1$

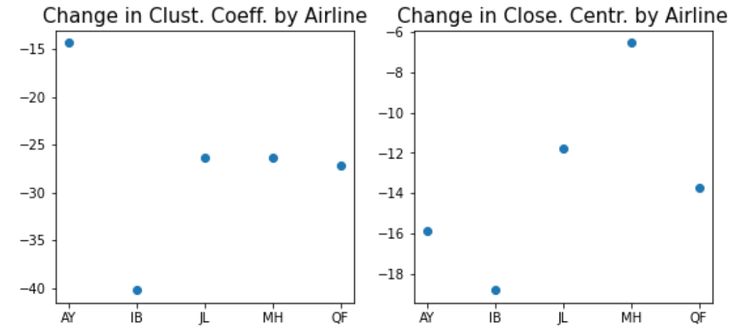
Result

One World

$c_1, c_2, c_3, c_4, c_5, c_6$ = (-0.001, -0.002, -0.133, 1.161, 0.005, -0.029)







Result

One World

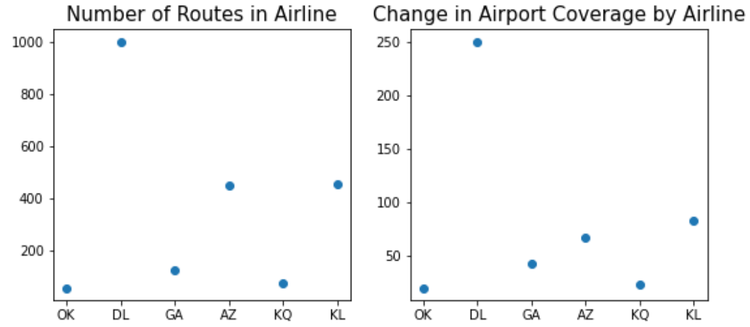
前五推薦加入：
代號：['7R' 'BE' 'IG' '9E' '9R']
名稱：['BRA-Transportes Aereos' 'Flybe' 'Meridiana' 'Pinnacle Airlines' 'SATENA']

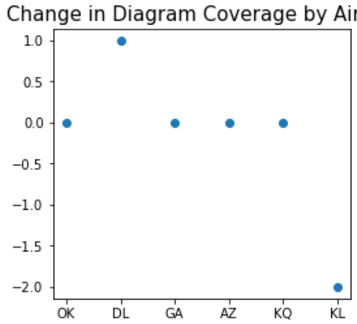
$c_1, c_2, c_3, c_4, c_5, c_6$ = (-0.001,-0.002,-0.133,1.161,0.005,-0.029)

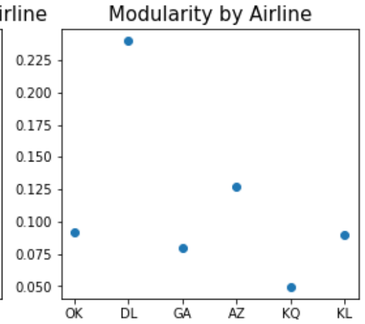
Result

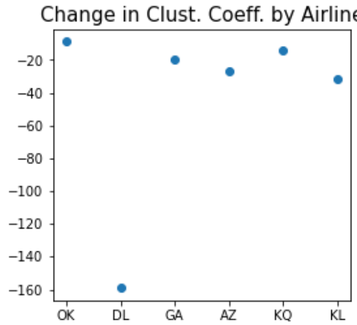
Skyteam

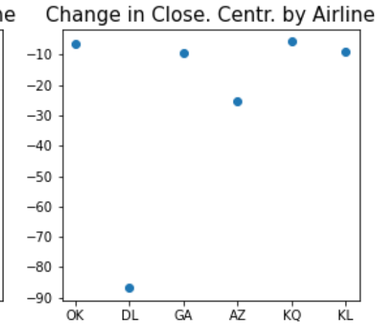
$c_1, c_2, c_3, c_4, c_5, c_6 = (0.001, -0.003, 0.148, 0.842, -0.005, 0.018)$











Result

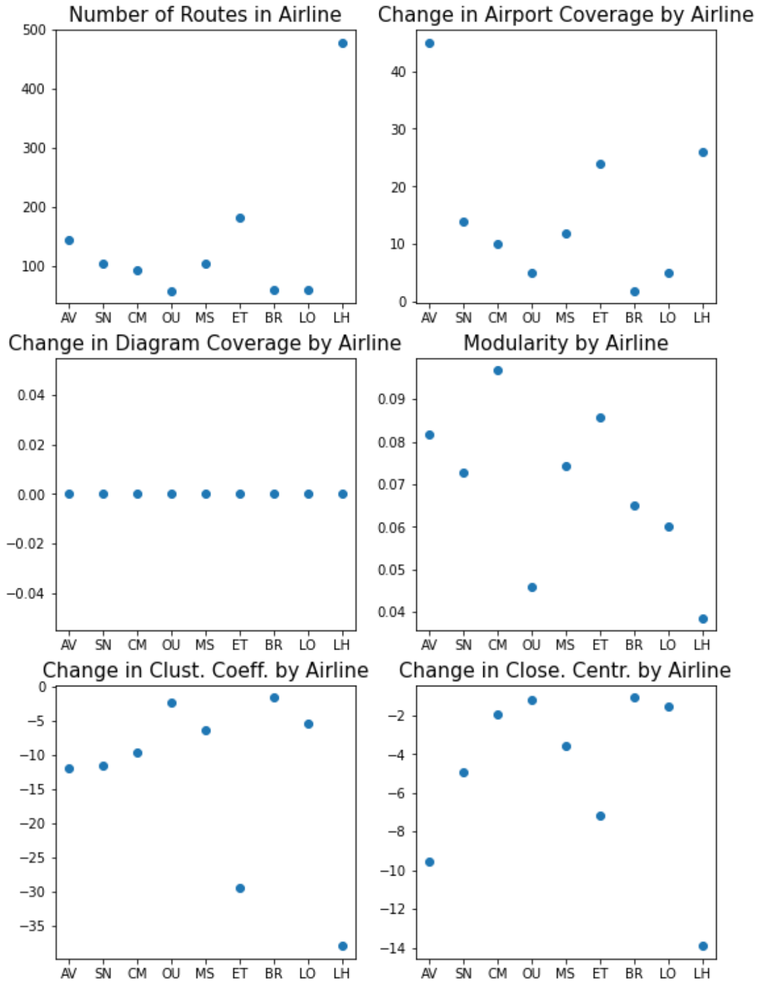
Skyteam

前五推薦加入：
代號：['AA' 'FR' 'US' 'UA' 'CZ']
名稱：['American Airlines' 'Ryanair' 'US Airways' 'United Airlines'
'China Southern Airlines']

Result

Star Alliance

$c_1, c_2, c_3, c_4, c_5, c_6 = (0.000, 0.000, 1.000, 0.000, -0.000, 0.000)$



Result

Star Alliance

前五推薦加入：
代號：['AB' 'GQ' 'A2' 'AD' 'FR']
名稱：['Air Berlin' 'Big Sky Airlines' 'Cielos Airlines' 'Azul' 'Ryanair']

Conclusion

- Each case produce the smallest magnitude coefficients

- Some alliance have similar features

- Each alliance had fewer than half of the tested airlines fail