Metody analýzy dat II

Zkoumaný dataset: síť letišť

Dostupné z <https://toreopsahl.com/datasets/#usairports>, resp. z <https://openflights.org/>

Informace o vrcholech dostupné z <http://opsahl.co.uk/tnet/datasets/openflights_airports.txt>

Síť reprezentuje lety mezi letištěmi po celém světě. Jednotlivé vrcholy jsou letiště a hrany lety mezi nimi. Síť je orientovaná a vážená.

# Základní informace

Počet vrcholů: 2939

Počet hran: 30501

Průměrný stupeň: 10,378

Průměrný vážený stupeň: 15,128

Průměrná vzdálenost: 4.097

Počet clusterů: 11

Velikosti clusterů: 2905, 7, 9, 2, 2, 2, 2, 4, 2, 2, 2

Nejdelší nejkratší cesta: 16

Modularita:

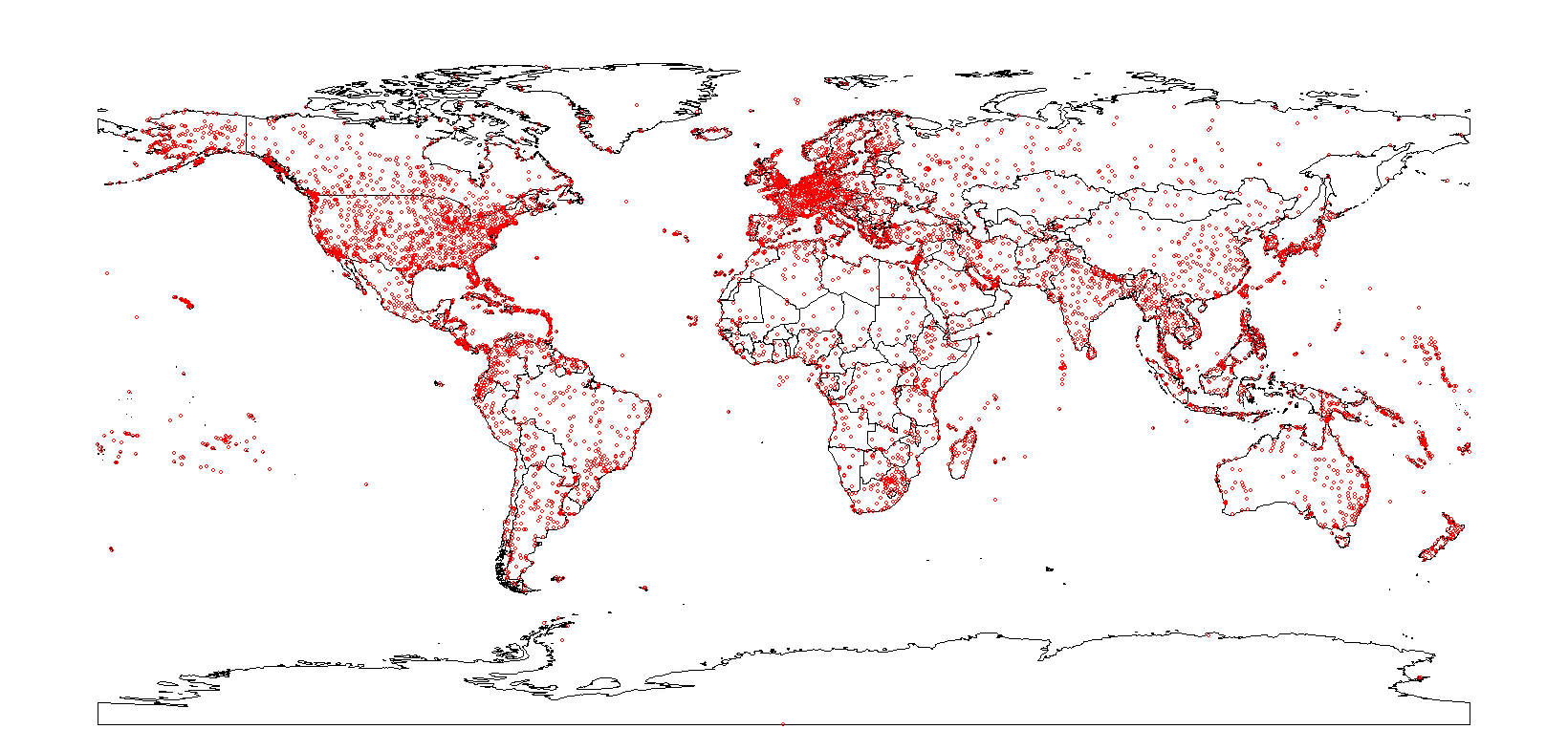
* Q: 0.669 až 0.674
* Počet komunit: 26 až 34
* Pro 34 komunit byla modularita 0.669, pro 26 komunit 0.674.
* R mi najde 35 komunit a modularitu 0.67

Nejfrektovanější letiště (podle stupně):

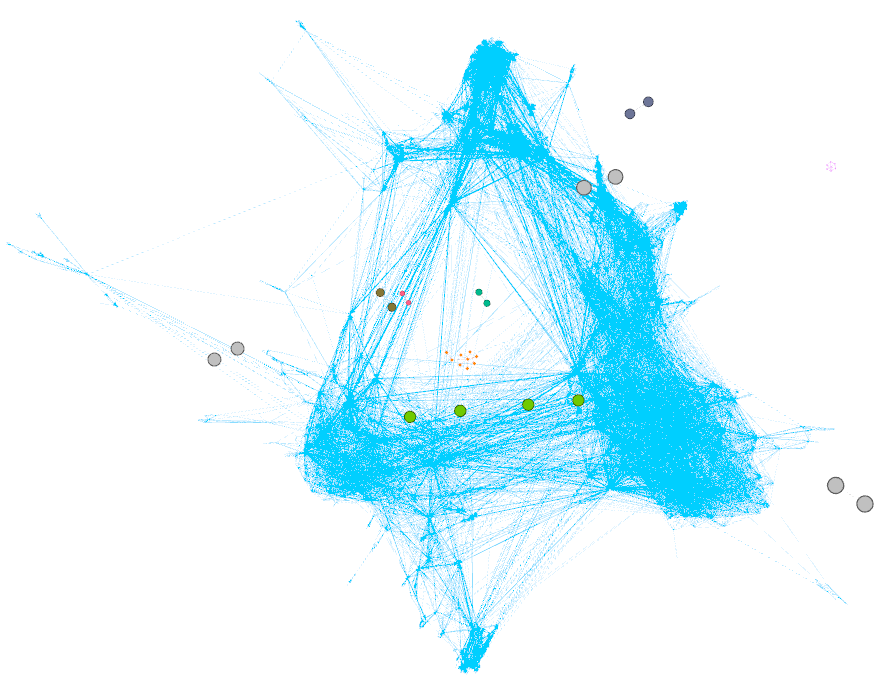
* Frankfurt Main (Frankfurt, Německo) (340)
* Charles De Gaulle (Paříž, Francie) (1382)
* Schiphol (Amsterdam, Nizozemsko) (580)

Nejdůležitější letiště (podle hubů (autority)):

* Frankfurt Main (Frankfurt, Německo) (340)
* Schiphol (Amsterdam, Nizozemsko) (580)
* Charles De Gaulle (Paříž, Francie) (1382)



Obrázek : Mapa letišť.

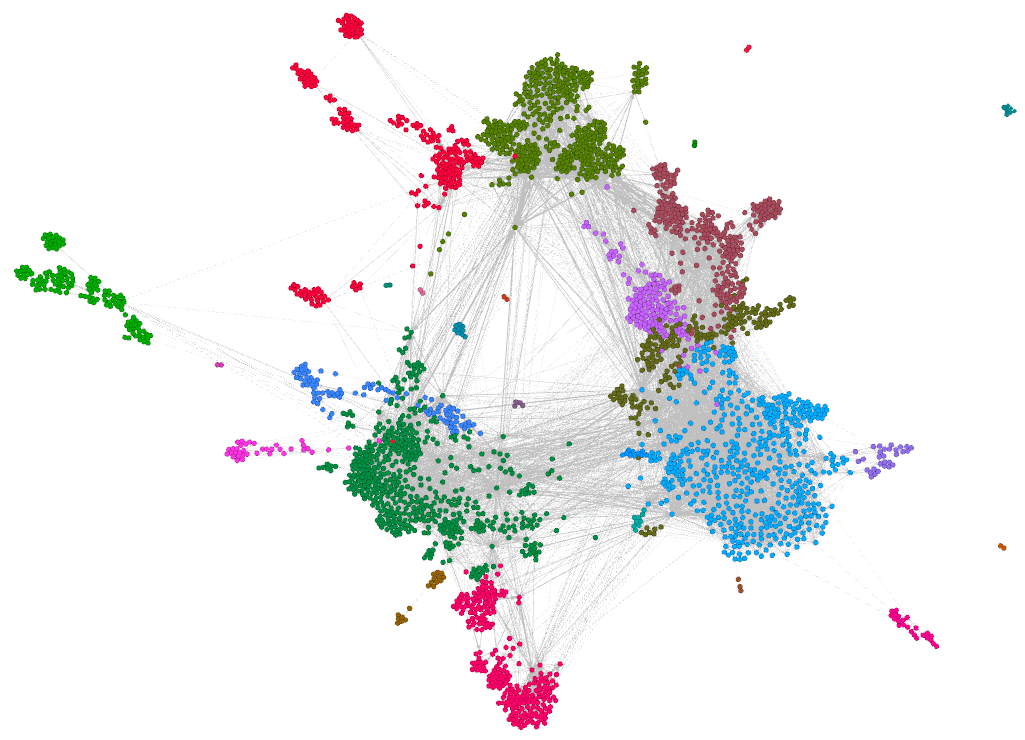


Obrázek : Shluky / mírně propojené komponenty.

Hlavní síť tvoří 2905 letišť, zbylé podsítě tvoří většinou ostrovy jako např.:

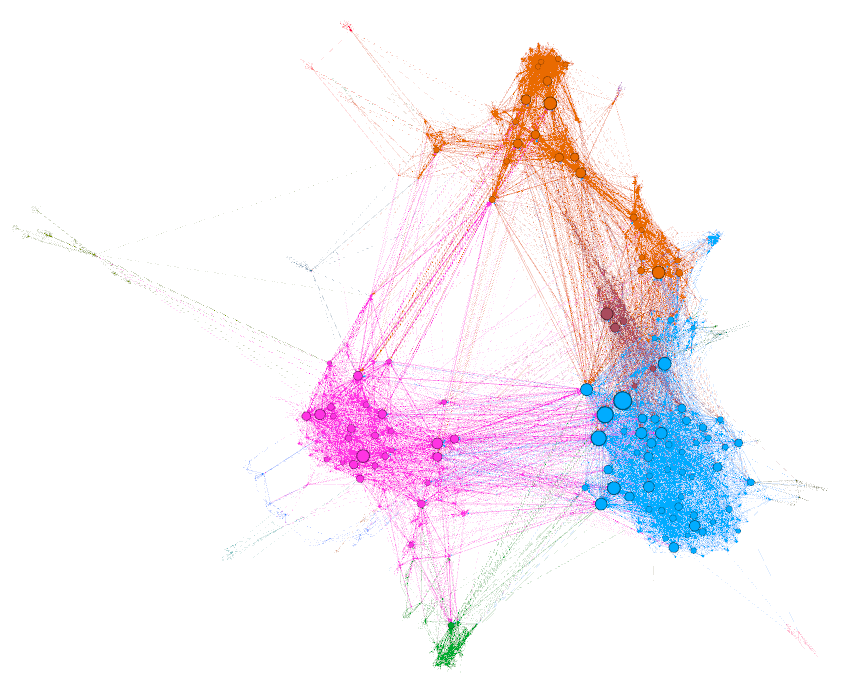
* 9 letišť na ostrově New Caledonia (Austrálie) (Ids -1998-2004, 5919-5921).
* 7 letišť na ostrovech u Irska (UK) (Ids - 529, 5564-5569, 5571).

Nebo spoje mezi velmi malými letišti: Spoj mezi Christiansted Harbor Seaplane Base (Christiansted, Virgin Islands) a Scappoose Industrial Airpark (Oregon, Portland).



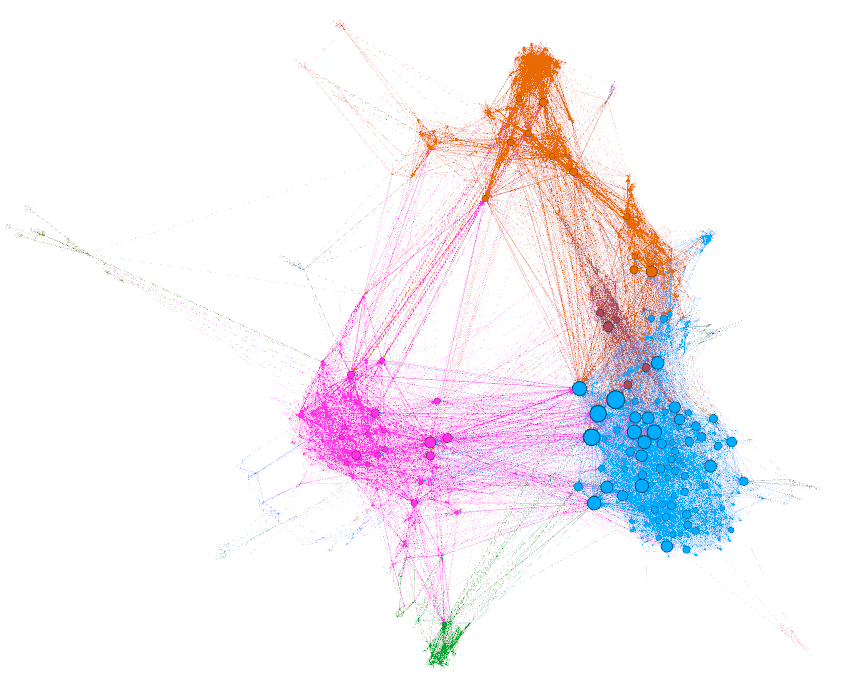
Obrázek : Detekce komunit (pomocí modularity).

Algoritmus pracuje s náhodou a detekoval komunity v rozmezí 26-34.



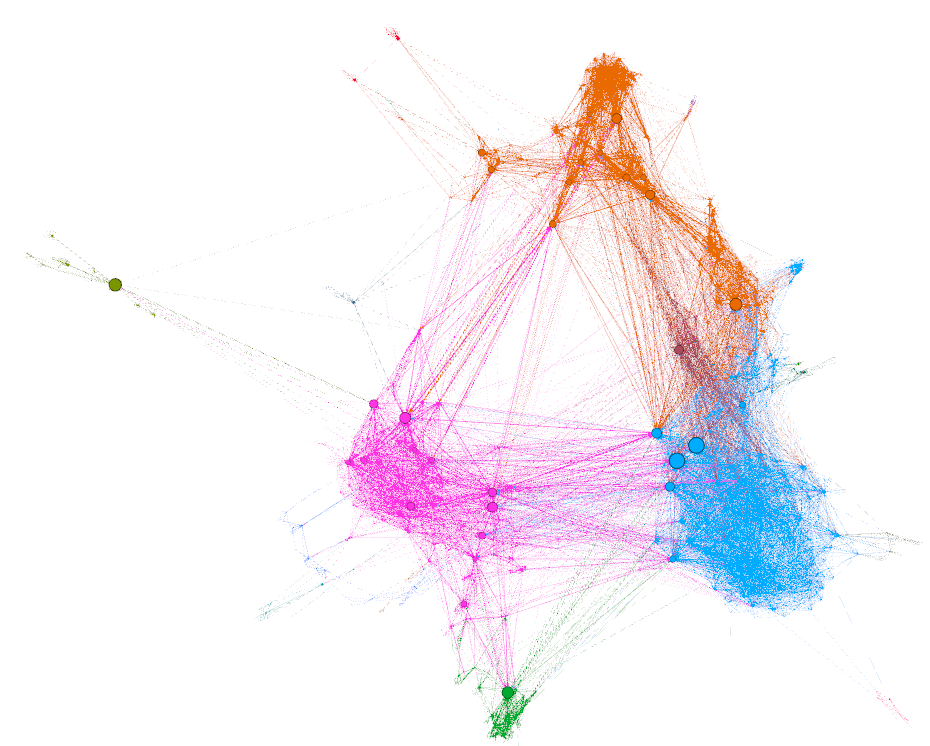
Obrázek : Velikost vrcholů reprezentuje počet cest k ostatním letištím (stupeň).

Jen 150 letišť je propojeno s více jak 100 dalšími letišti.



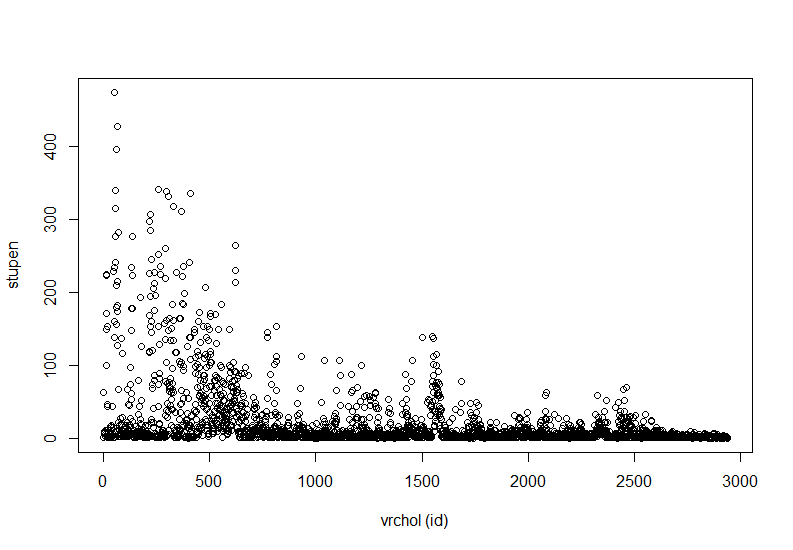
Obrázek : Velikost vrcholů repzetentuje jak moc je letiště hub.

Obdobné jako obrázek předtím, ale více zvýrazní huby v jednotlivých komunitách, stupeň je globálně ke všem.

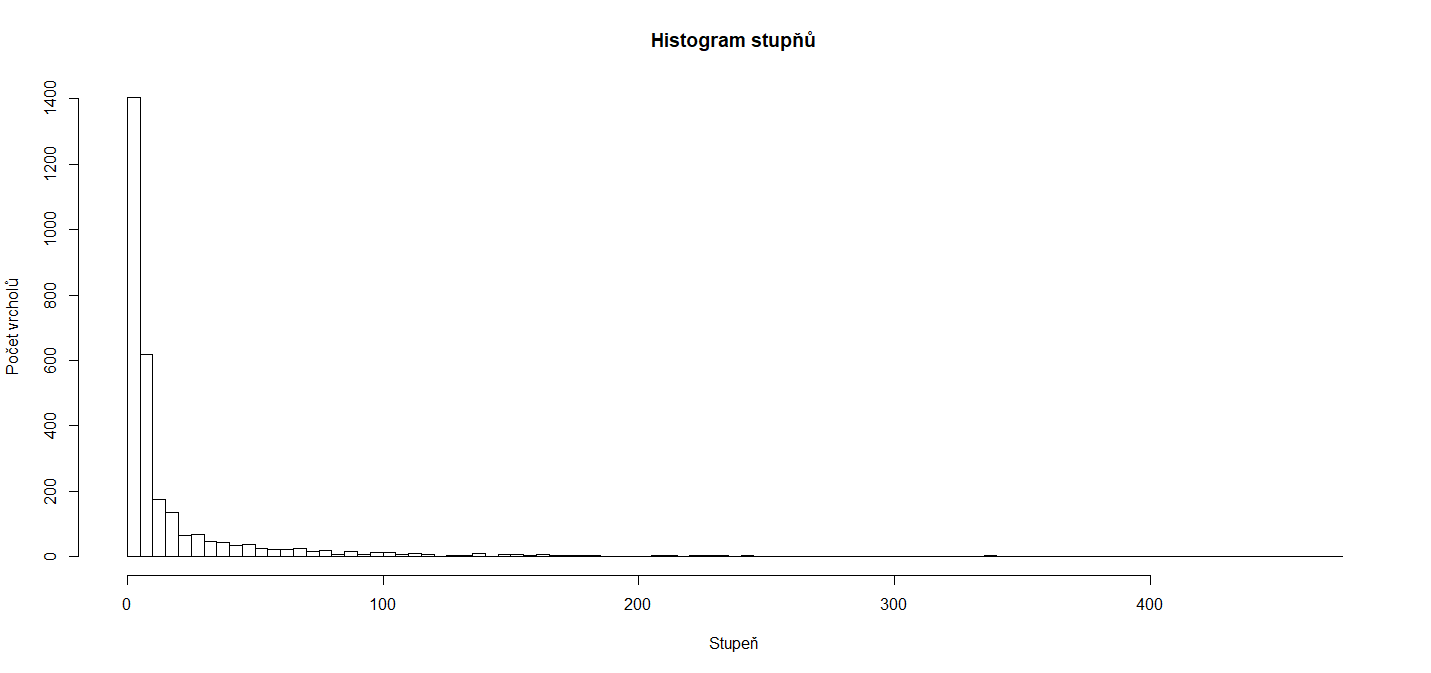


Obrázek : Velikost vrcholů reprezentuje betweeness centralitu (počet nejkratších cest skrz letiště).

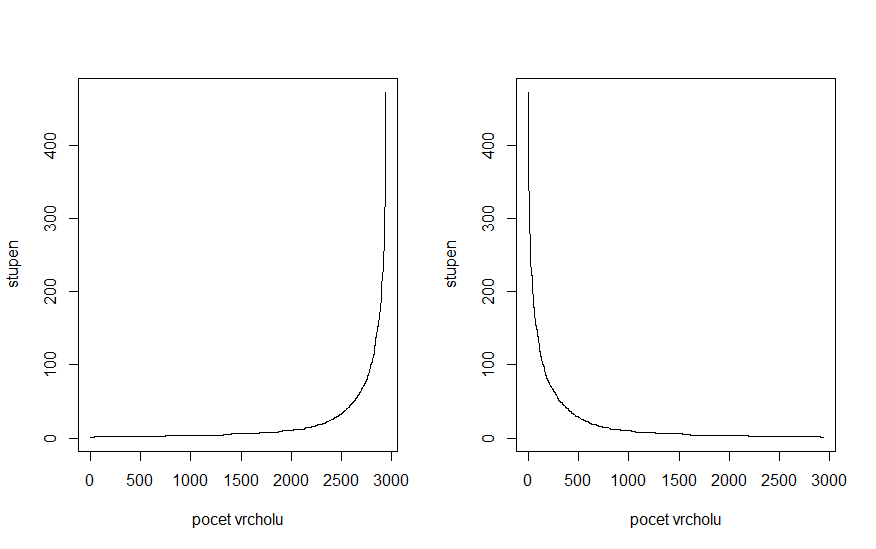
Jde vidět, že jednotlivé komunity mají obvykle jedno centrální letiště odkud se létá všude dál (zelená vlevo nahoře a uprostřed dole).



Obrázek : Stupeň jednotlivých vrcholů.

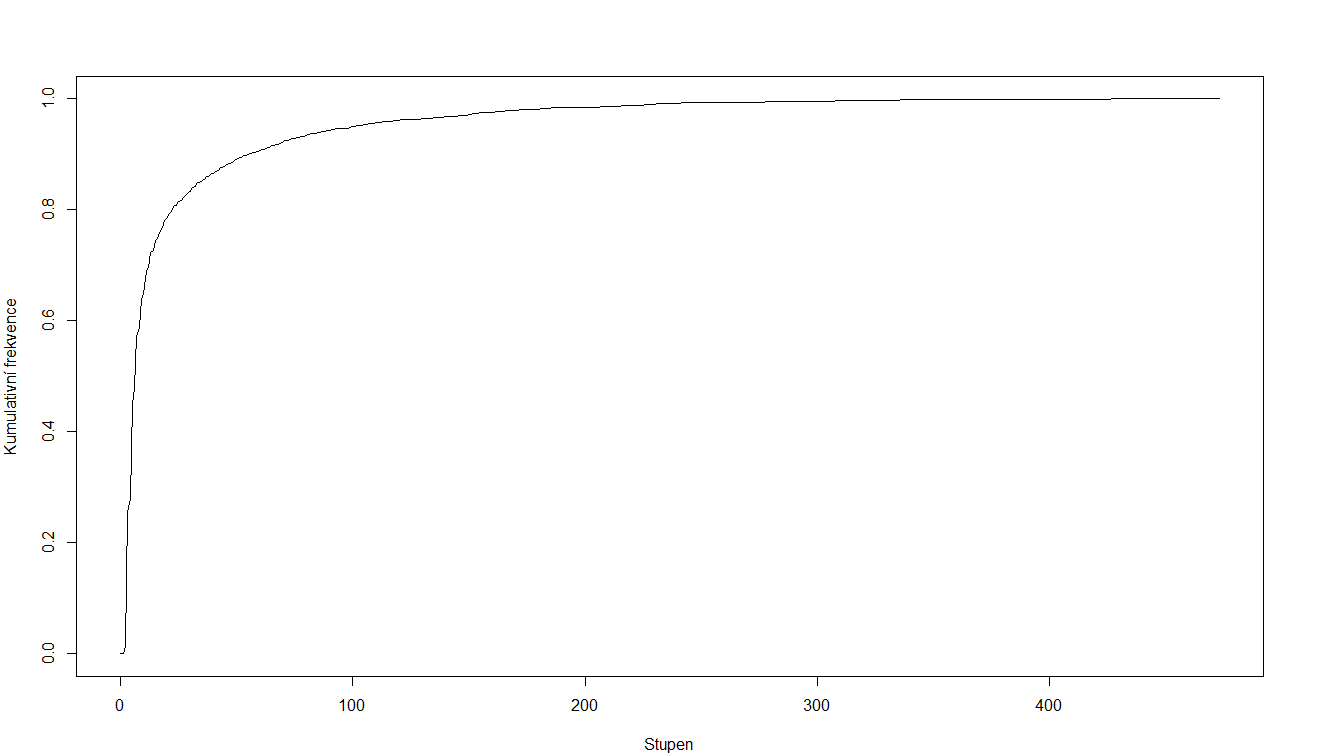


Obrázek : Histogram stupňů.



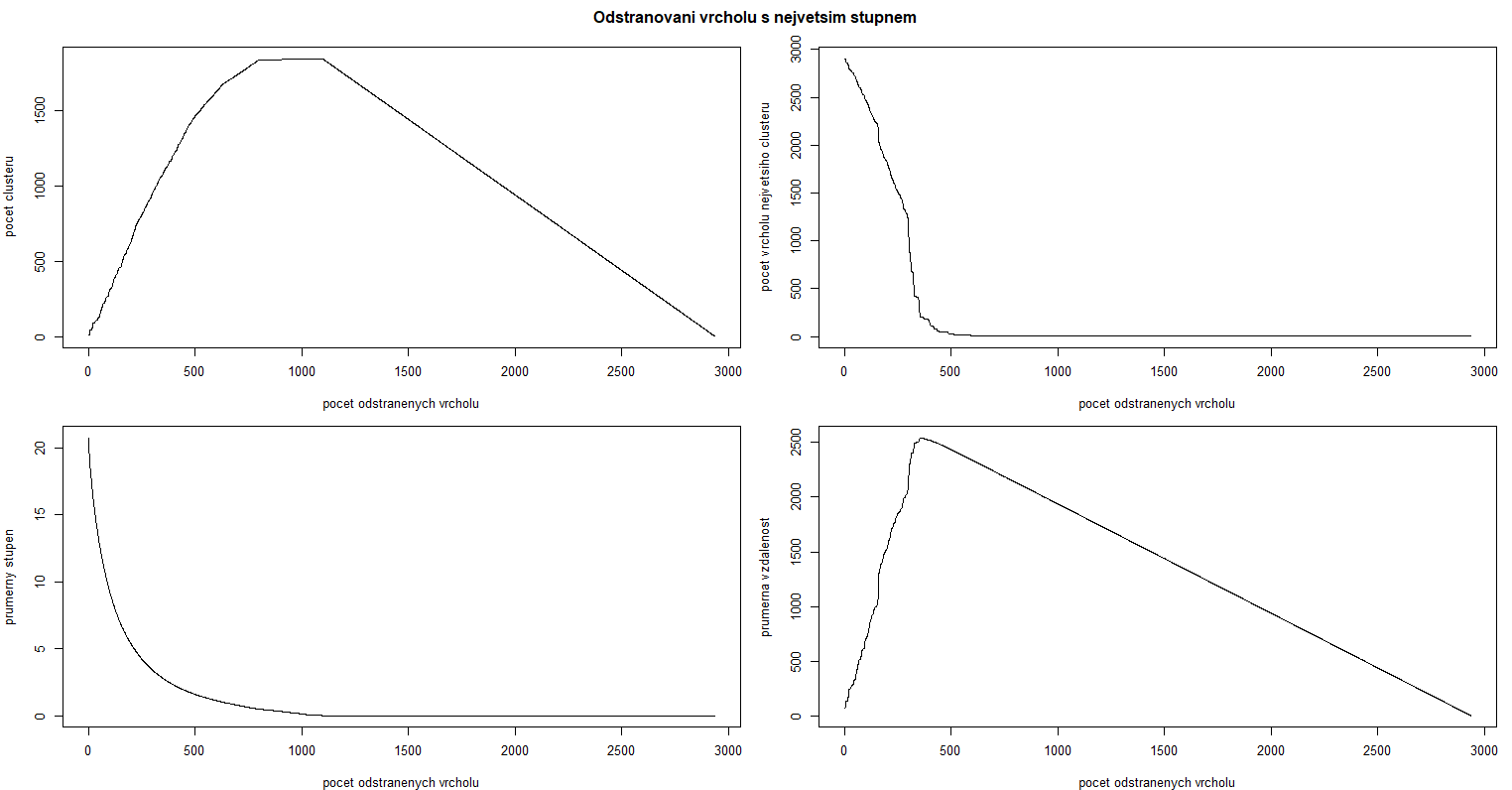
Obrázek : Distribuce stupňů vrcholů (seřazeno).

Jen 150 letišť na světě lítá do více jak 100 ostatních letišť.

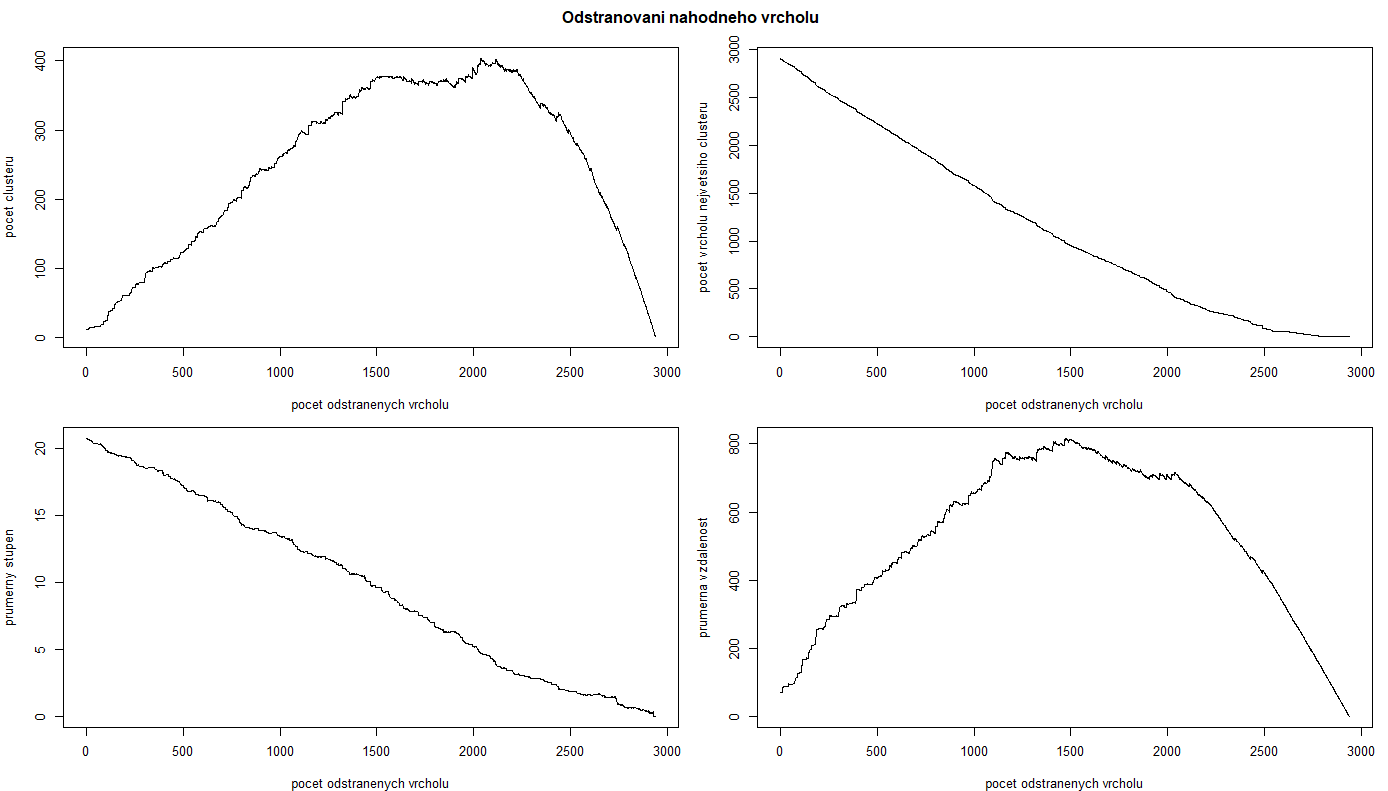


Obrázek : Kumulované distribuce stupně.

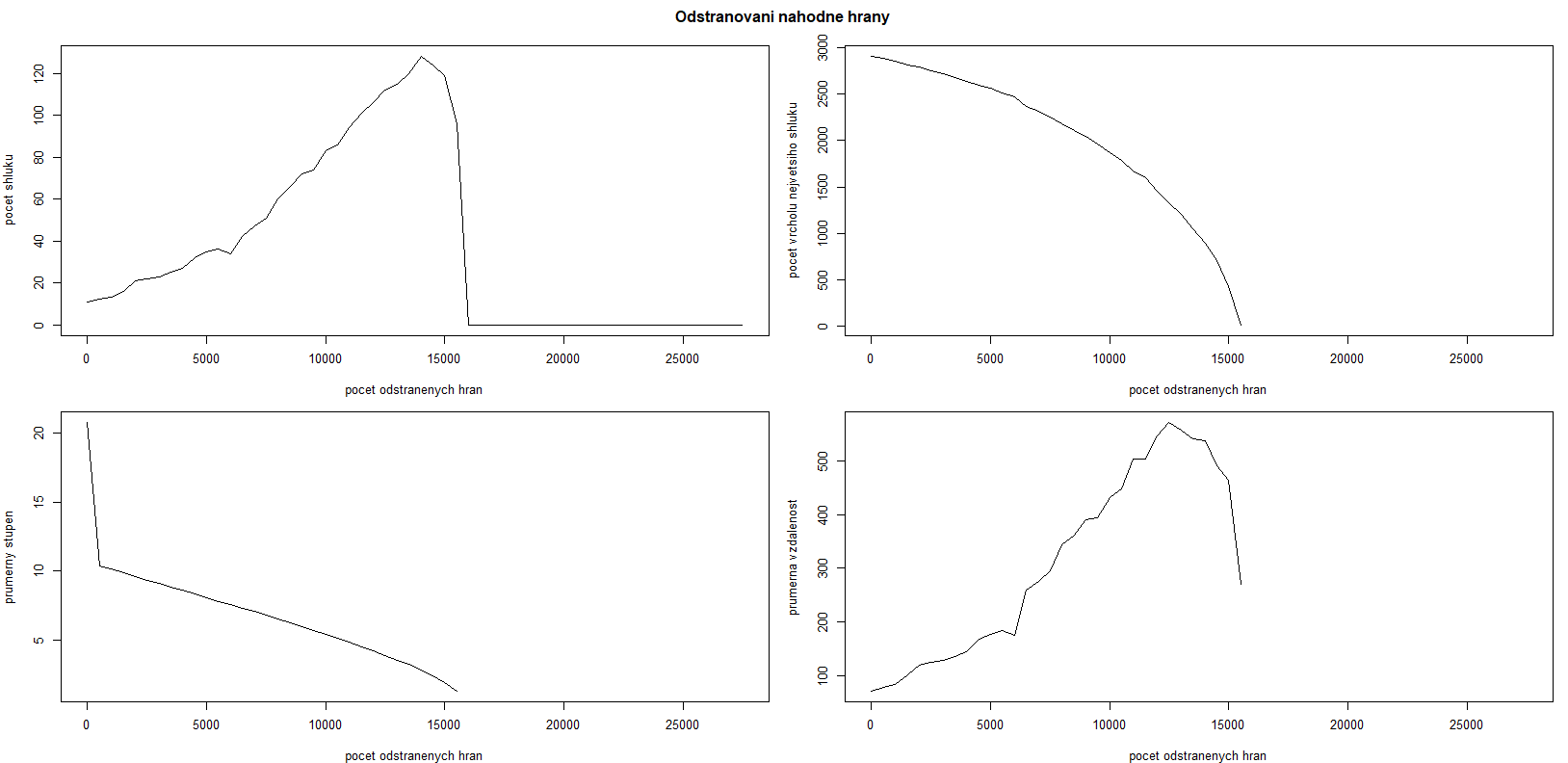
# Odolnost sítě



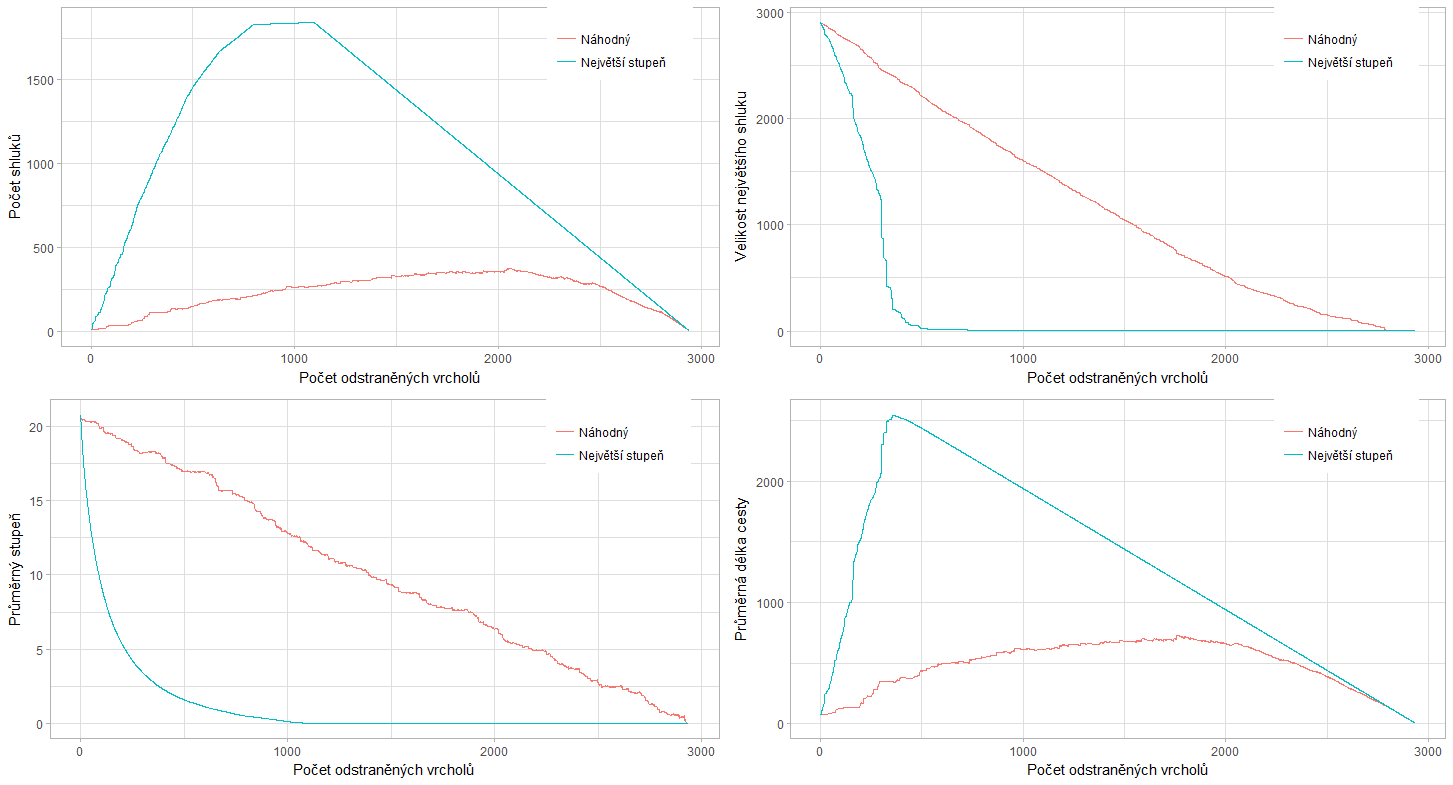
Obrázek : Odolnost síťě - odebíráme vrchol s největším stupněm.



Obrázek : Odolnost sítě – odebírání náhodného vrcholu.



Obrázek : Odolnost sítě – odebírání náhodné hrany (vizualizováno co 500 odebraných hran).



Obrázek : Odolnost sítě – porovnání odstranění náhodného vrcholu a vrcholu s největším stupněm.