

# Pajek

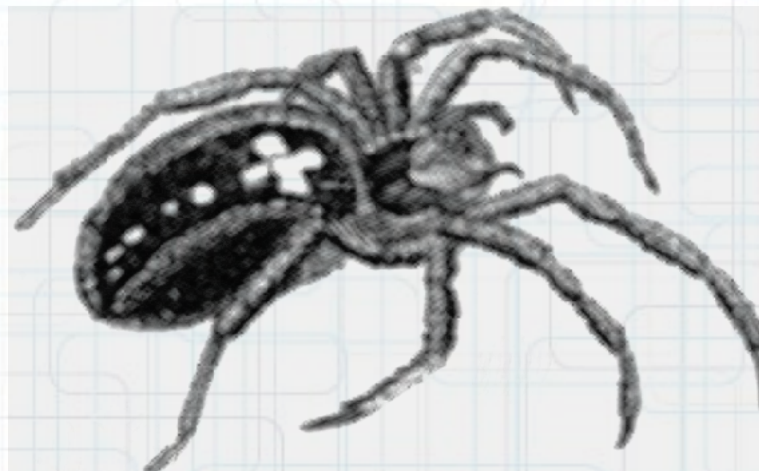
## PV115 Laboratoř dobývání znalostí

Masarykova univerzita, fakulta informatiky

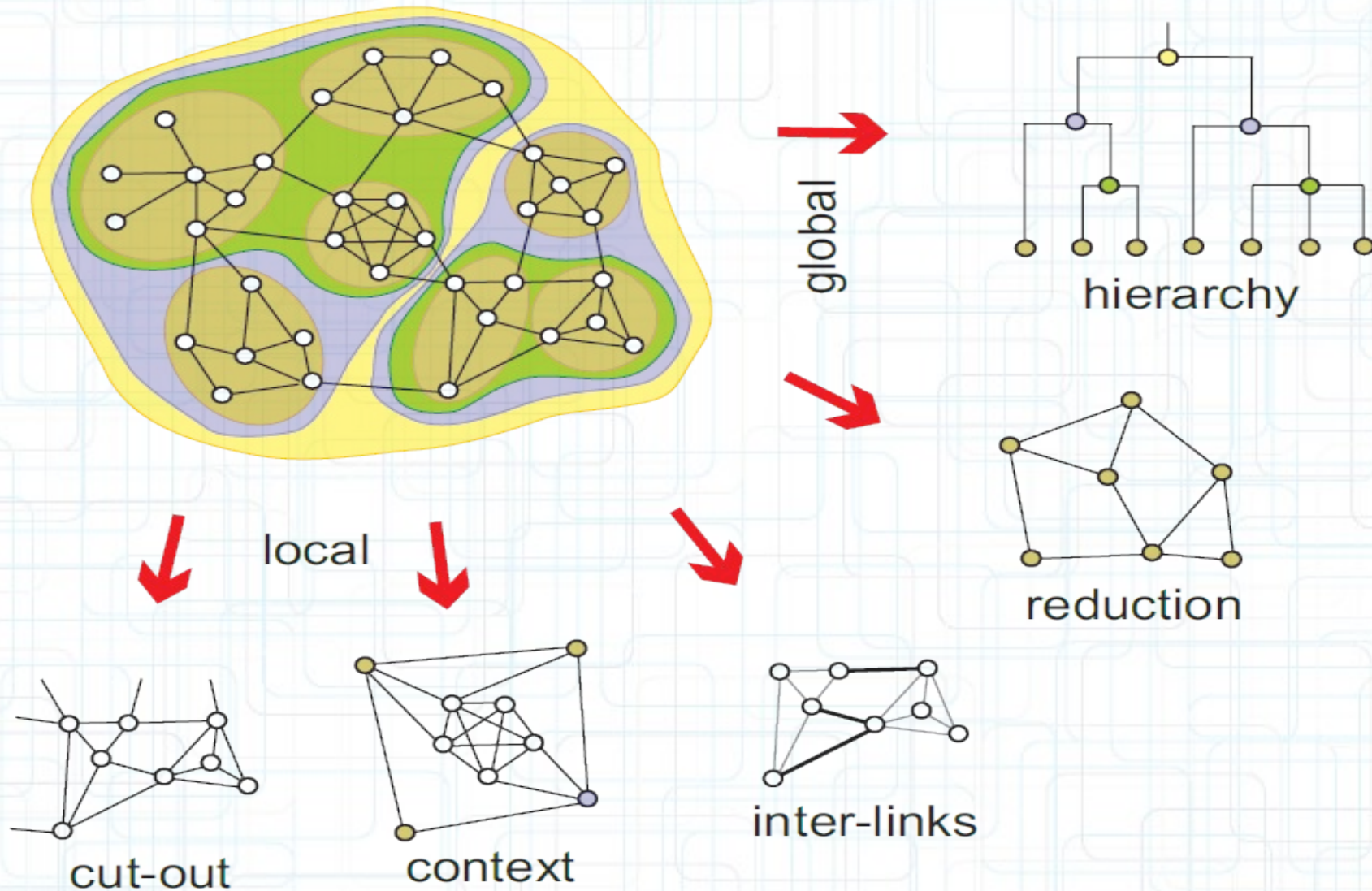
Juraj Jurčo, 173001@mail.muni.cz

# Pajek

- Slovinský program, vývoj od roku 1996
- V slovinčine pajek znamená pavúk
- Program na analýzu a vizualizáciu (veľkých) sietí
  - molekuly v chémii, interakčné siete proteínov, internetovej siete, citačné siete, dolovanie dát...
  - príklady veľkých sietí: <http://vlado.fmf.uni-lj.si/pub/networks/data/>



# Prístup práce s veľkými sieťami



# Ciele programu Pajek

- Uľahčiť predstavivosť pomocou (rekurzívnej) dekompozície veľkých sietí na menšie
- Poskytnúť užívateľovi výkonné vizualizačné metódy
- Implementácia vybraných efektívnych algoritmov na analýzu veľkých sietí

# Členenie programu

- **Networks**
  - Siete, s ktorými pracujeme, prípadne už upravené siete
- **Partitions**
  - Ktorý vrchol ku ktorej skupine vrcholov patrí
- **Vectors**
  - Každý vrchol môže obsahovať niekoľko číselných atribútov
- **Permutations**
  - Preusporiadanie vrcholov

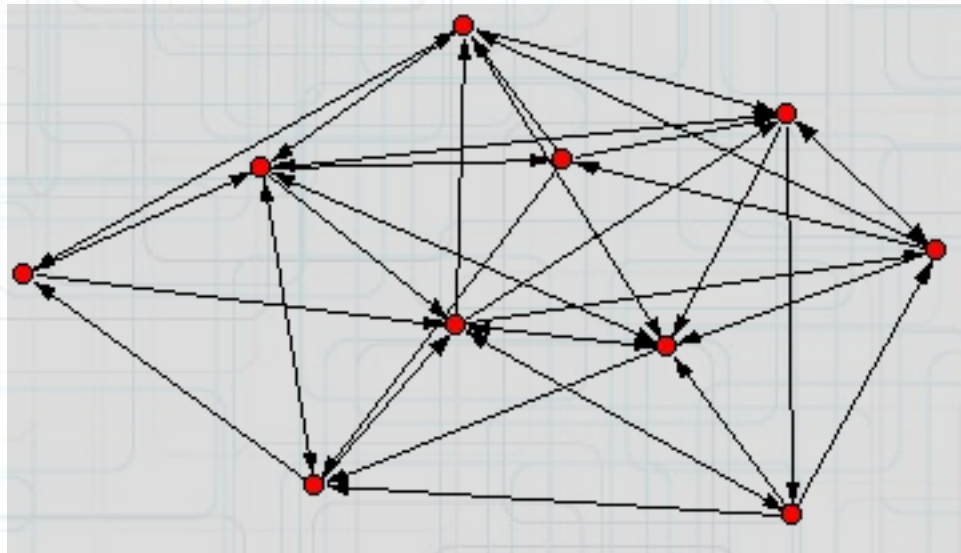
# Členenie programu

- Cluster
  - Podmnožina vrcholov
- Hierarchy
  - Hierarchicky usporiadané vrcholy

Root  
g1 g2  
g11 g12 v5,v6,v7  
v1,v2 v3,v4

# Draw

- Vykresľovanie grafu, partície alebo vektoru
- Automatické usporiadanie grafu
- Otáčanie, zoom, ofarbenie, zobrazenie popiskov



# Net

- Transformácia siete

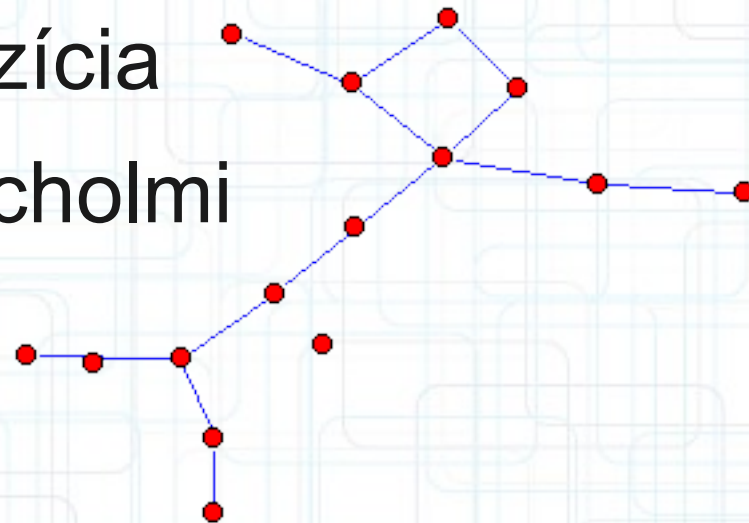
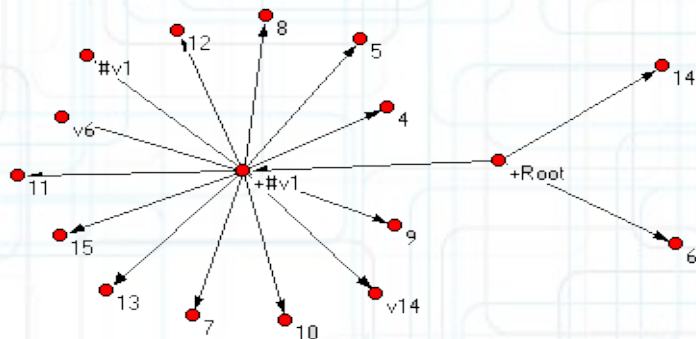
- Pridanie/odobratie/úprava uzlov/hrán
- Redukcia grafu – odstránenie vrcholov a hrán podľa stupňa
- Generovanie sietí v čase

- Náhodne generované siete

- Jednoduché, Erdos-Renyi, acyklické, (ne)orientované...

- Hierarchická dekompozícia

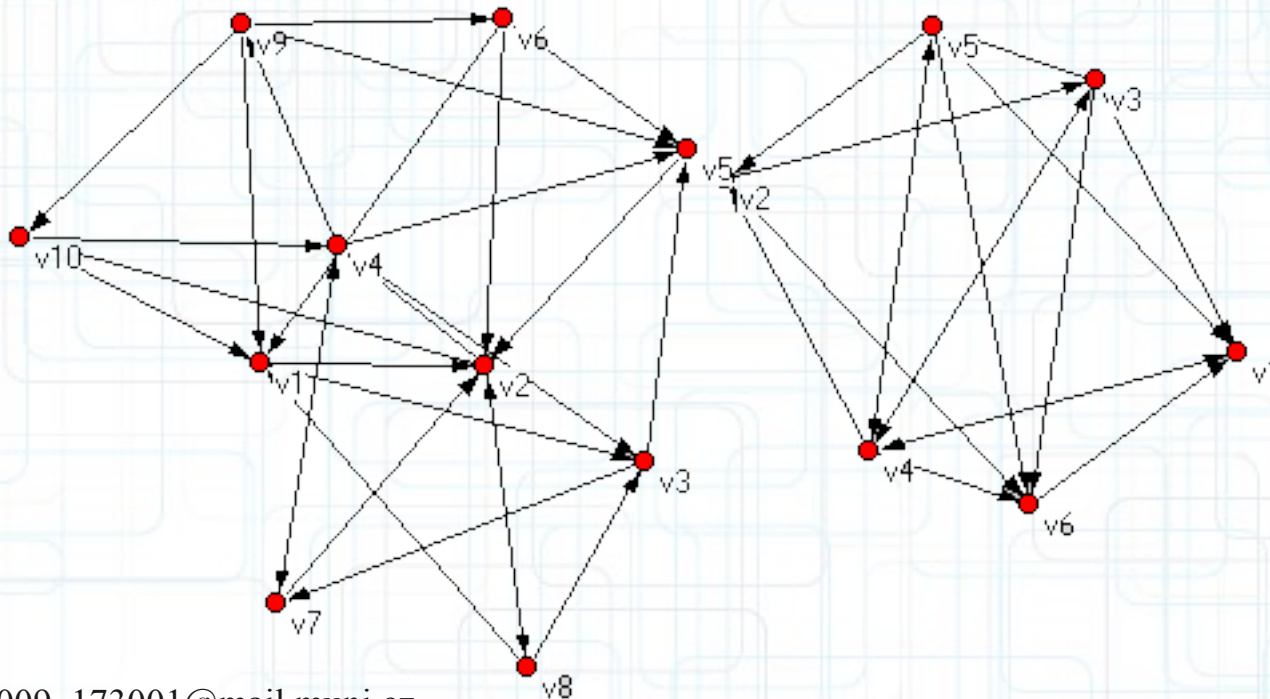
- Cesta medzi dvoma vrcholmi





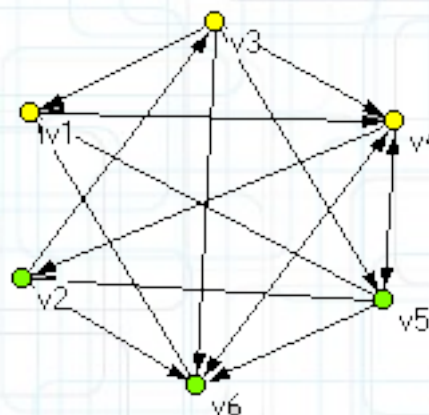
# Nets

- Zjednotenie riadkov (sietí)
- Prienik
- Rozdiel
- Hľadanie fragmentov jednej siete, v inej



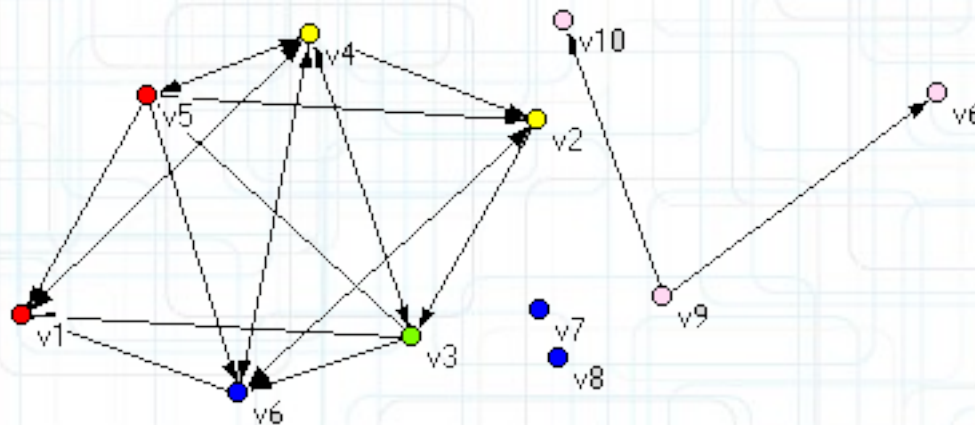
# Partition

- Vytvorenie konštantnej partície siete
  - Všetky vrcholy v jednej partícii
- Vytvorenie náhodnej partície
  - Vrcholy sú náhodne rozdelené
- Kanonická partícia
- Transformácia partície
  - do siete, permutácie, klastru, vektoru, hierarchie



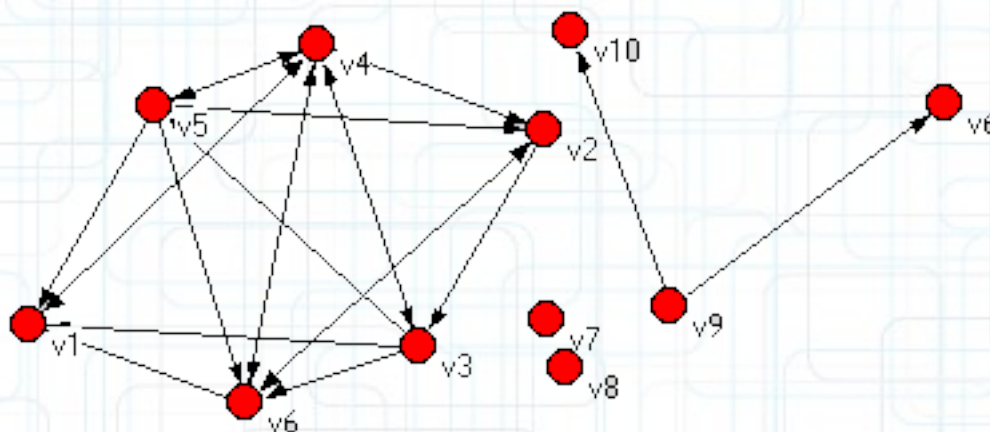
# Partitions

- Minimum/maximum dvoch partícií
- Spojenie dvoch partícií
- Prienik



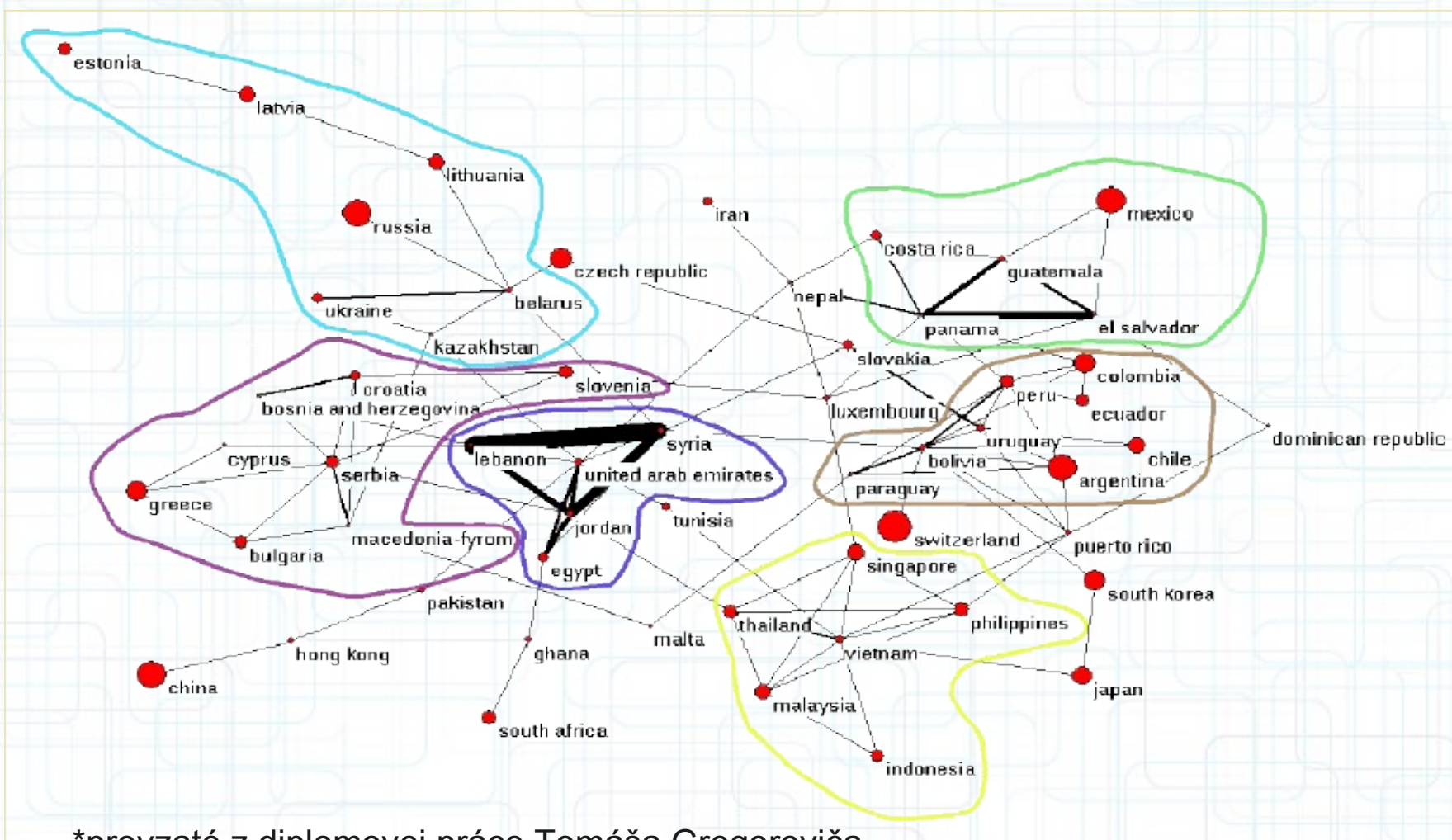
# Vector

- Vytvorenie konštantného vektora
- Extrakcia podvektora z daného vektora
- Vytvorenie partície/permutácie z vektora
- Transformácia vektora pomocou funkcie
  - Ln, exp, mocnina, pre násobenie konštantou, normalizácia, invertovanie hodnôt...



# Vektor - príklad

- Aktéri grafu zlúčení do uzlov podľa krajiny



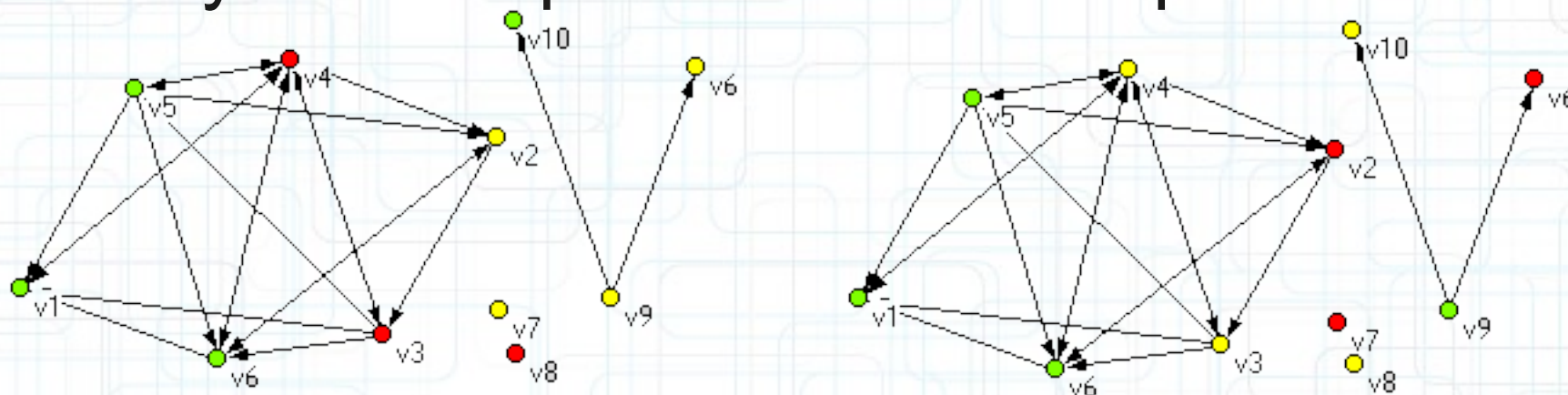
\*prevzaté z diplomovej práce Tomáša Gregoroviča

# Vectors

- Sčítanie/odčítanie vektorov
- Násobenie/delenie
- Lineárna regresia
- Minimum/maximum
- Transformácia
  - Karteziánska  $\leftrightarrow$  polárna

# Permutation

- Náhodná
- Identita
- Inverzná
- Zrkadlovo obrátená
- Vytvorenie partície/vektoru z permutácie



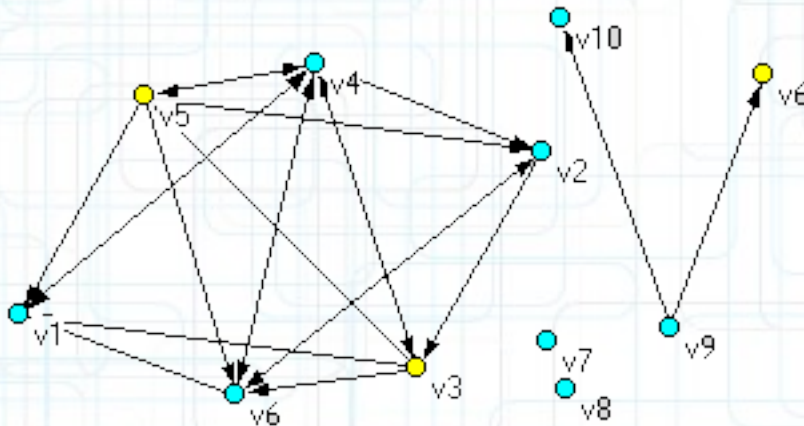
# Permutations

- Fúzia permutácií
  - Pridanie druhej permutácie na koniec prvej



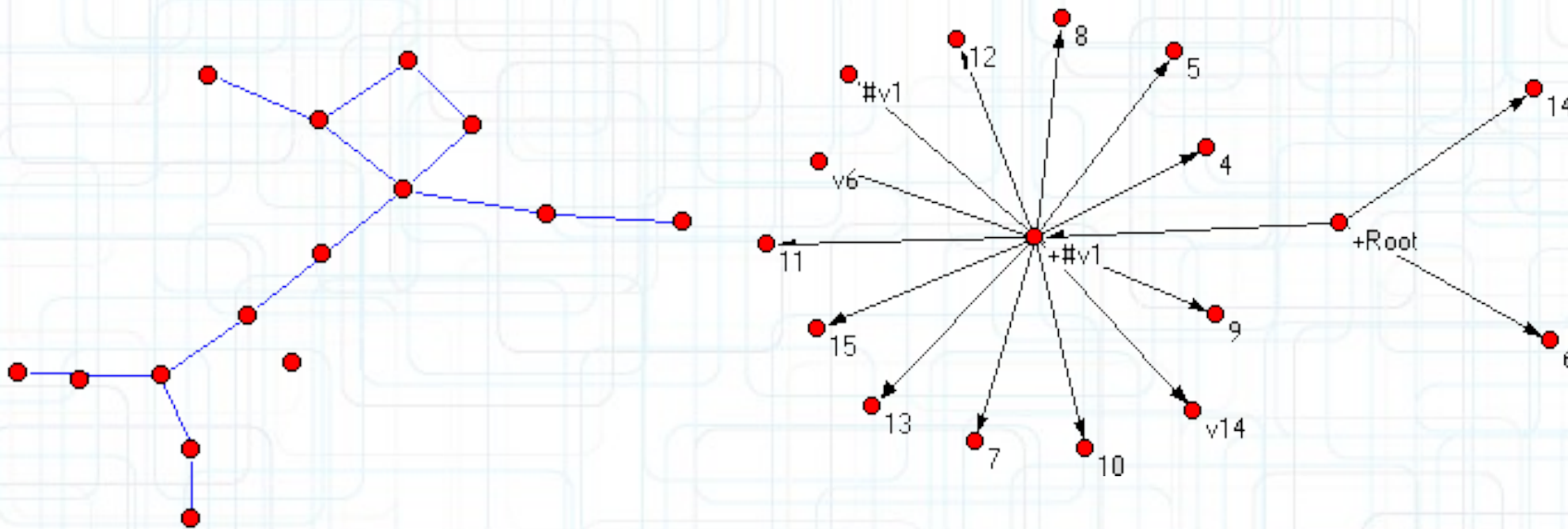
# Cluster

- Vytvorenie prázdneho/kompletného klastra
- Vytvorenie partície z klastra



# Hierarchy

- Extrakcia klastra
- Vytvorenie siete/permutácie/klastra z hierarchie



# Zhrnutie

- Program zvláda aj grafy s miliónmi vrcholov (3.6M) pomerne rýchlo, záleží však aj od aplikovanej funkcie
- Program zameraný najmä na funkcionálnosť
- Funkcionálnosť je dobre popísaná v manuáli programu
- Program úspešne použil Tomáš Gregorovič pri analýze sociálnej siete *CouchSurfing* s viac ako 370000 užívateľskými profilmi.
  - [https://is.muni.cz/th/139855/fi\\_m/dp.pdf](https://is.muni.cz/th/139855/fi_m/dp.pdf)

# Literatúra

- Batagelj, Vladimir; Mrvar, Andrej. Pajek [počítačový program]. Ver. 1.26. [Slovinsko], 1998 [citované 2010-02-06]. Dostupné z <<http://vlado.fmf.uni-lj.si/pub/networks/pajek/pajek126.exe>>. Zdarma pre nekomerčné použitie.
- Batagelj, Vladimir; Mrvar, Andrej. Pajek - Program for Analysis and Visualization of Large Networks: Reference Manual, 5. Máj 2009, posledná aktualizácia 7. Január 2010.  
URL: <<http://pajek.imfm.si/lib/exe/fetch.php?id=download&cache=cache&media=dl:pajekman126.pdf>>.
- Boldiš, Petr. *Bibliografické citace dokumentů podle CSN ISO 690 a CSN ISO 690-2: Část 2 – Modely a příklady citací u jednotlivých typů dokumentů*. Verze 3.0 (2004). c 1999–2004, posledná aktualizácia 11. 11. 2004.  
URL: <<http://www.boldis.cz/citace/citace2.pdf>>.
- Gregorovič, Tomáš. Extrakce informací ze sociálních médií, diplomová práca, 8. Február 2010.  
URL: <[https://is.muni.cz/th/139855/fi\\_m/dp.pdf](https://is.muni.cz/th/139855/fi_m/dp.pdf)>.