

Do more people say that they found a new job opportunity through a close friend or through an acquaintance?

Friendship graph

O PERSON





Onnela, J-P., Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L. Barabási. "Structure and tie strengths in mobile communication networks." *Proceedings of the national academy of sciences* 104, no. 18 (2007): 7332-7336.

https://www.pnas.org/doi/10.1073/pnas.0610245104



Image from: https://en.wikipedia.org/wiki/Community_structure



Zachary's Karate Club

Image from:

https://commons.wikimedia.org/wiki/File:Zachary%27s_karate_club.png

Lu, Zhenqi, Johan Wahlström, and Arye Nehorai. "Community detection in complex networks via clique conductance." *Scientific reports* 8, no. 1 (2018): 1-16.

https://www.nature.com/articles/s41598-018-23932-z



https://bigdata.oden.utexas.edu/project/graph-clustering/



Infomap

How it works

Pick a starting edge
At edge, randomly choose another edge and keep going
(stop on some pre-defined condition)
Repeat this process many times

=> Across whole dataset, will spend more time following edges inside communities than moving between them

Rosvall, Martin, and Carl T. Bergstrom. "Maps of random walks on complex networks reveal community structure." *Proceedings of the national academy of sciences* 105, no. 4 (2008): 1118-1123.

https://www.mapequation.org/assets/publicat ions/RosvallBergstromPNAS2008Full.pdf



ET. ONE WALK: ABCDEFEFGH...

Problem: will be doing lots and lots of walks Node LinkedList<Node>



These walks are going to get long, especially if going randomly

=> So how can we store the walk more efficiently?

Infomap

Insight: use compression, creative bit encodings for nodes to save space

Consider a single Node object we might use to represent a graph:





Since a walk will involve nodes that are close together, use this to figure out bit encoding?

Infomap





 $\begin{array}{c} \underline{111}\\ 0000 & 11 & 01 & 101 & 100 & 101 & 01 & 0001 \\ \hline 11 & 000 & 11 & 000 & 111 & 10 & 011 & 000 & 111 & 10 & 010 & 10 & 010 \\ \hline 11 & 000 & 111 & 0001 & \underline{0} & 111 & 010 & 010 & 011 & 00 & 011 & 00 & 111 & 00 & 111 \\ \hline 10 & 111 & 110 & 1011 & \underline{111} & 01 & 101 & 010 & 0001 & \underline{0} & 110 & 011 & 100 & 111 & 0111 \\ \hline 10 & 111 & 000 & 10 & 0001 & \underline{0} & 111 & 010 & 100 & 010 & 011 & 1011 \\ \hline 10 & 111 & 000 & 10 & 0001 & \underline{0} & 111 & 010 & 1010 & 010 & 1011 & 1011 \\ \hline 10 & 111 & 000 & 10 & 0001 & \underline{0} & 111 & 010 & 1010 & 010 & 1011 & 111 \\ \hline 10 & 111 & 000 & 10 & 0001 & \underline{0} & 111 & 010 & 1010 & 010 & 1011 & 1011 \\ \hline \end{array}$