**IM problem**

networkx vs other packages

<https://www.timlrx.com/2019/05/05/benchmark-of-popular-graph-network-packages/>

# installing networkx

pip install networkx

pip install pytest

* **Graph realization**
  + **Node**: doc\_id

attributes: {query\_id: rel\_1d}

* **Edge**

relation between nodes

Q(q) - all queries that node1 and node2 participate

weight: ∑(q)/n

* **Implementation:**
  + **version 1:** 
    - step 1: create nodes
    - step 2: between every two nodes calculate weight
    - step 3: connect those nodes
    - **Conclusion:** takes several hours to be done (around 7 hours)
  + **version 2:**
    - step 1: create nodes
    - step 2: make list of all edges (in networkx edges can be assigned as list of tuples [(node1, node2)]
    - step 3: connect them at once
    - **Conclusion:** Failed in allocating memory for that many edges
* **How to run**
  + for creating a graph:

python generate\_graph.py

* + for loading a graph:

python generate\_graph.py load