Lecture 3 - Network Measures

Wednesday, January 17, 2024 2:49 PM

American Dialect Society's Word of the year for 2023: enshittification

Enshittification: aka 'platform decay', is the pattern of decreasing quality of online platforms that act as two- or multi-sided markets. E.g., Amazon, Google, Facebook

Step 1: Create a useful product to attract users (market 1).
Step 2: Monetize those users to sell ads (market 2).
Step 3: Get greedy, hurt market 1 to financially benefit from market 2.
Step 4: Start losing users, hurt market 2 to retain market 1.
Step 5: Basically, piss everyone off and destroy your brand.
Step 6: Die.

Review of last class

k-connectivity
- related to network resilience
→ how many vertices to remove to disconnect the graph

directed graphs
- edges are directed

$\circ \to \circ$

→ strong/weak connectivity

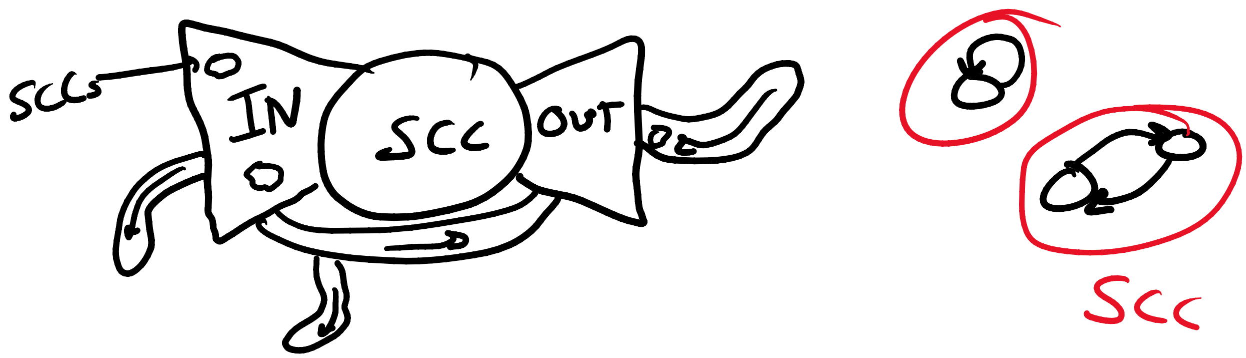
strong: $\forall u,v \in V(D): \exists\ u,v$-path

$\overset{u}{\circ} \to \circ \to \circ \to \overset{v}{\circ}$

weak: same as undirected connectivity if we ignore

connectivity if we ignore
direction of edge

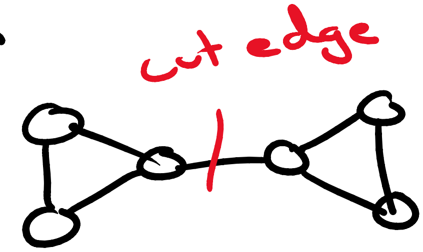
connectivity in general gives
insight into some global or
underlying structure of some graph



---

k-edge-connectivity

↳ how many edges to disconnect
a graph



Algorithmic approaches

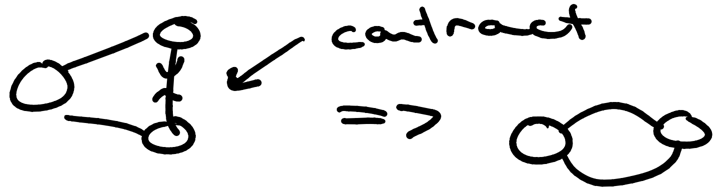
Network Flow $O(n^4)$

Faster algorithms

Random algos $\sim O(n^2)$

Random algos ~ $O(n^2)$

edge contraction

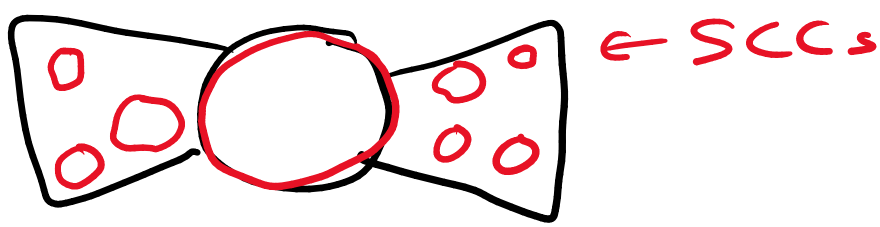


$\underbrace{\text{auxiliary graph}}_{\text{transformation using contraction or other technique}}$ construction

---

More on strong connectivity

context: strong connectivity decompositions

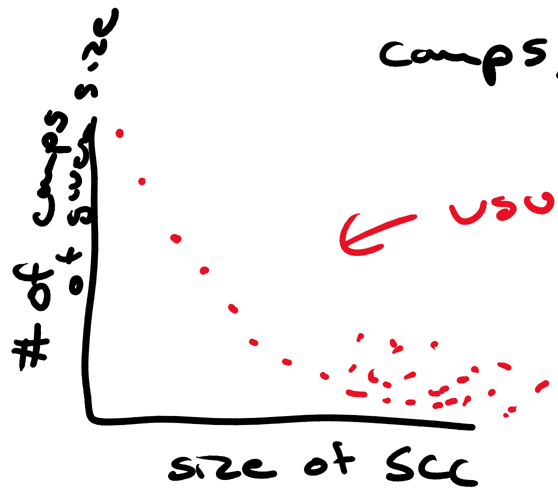


In terms of network measures:

size distributions of strong components

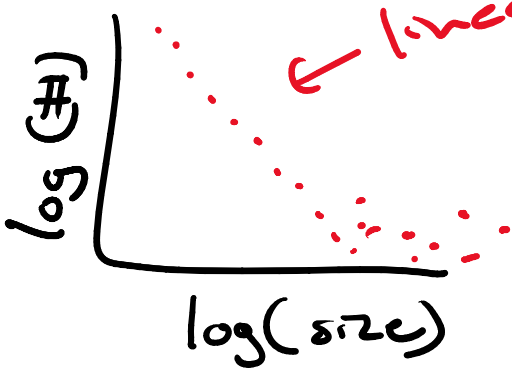
↳ how many size-1 strong comps, size-2, size-3 ... etc.

comps, size-2, size-3... etc.



← usually exponential decrease

← fat tail



← linear fit → power-law

← fat tail = more large SCCs than expected

aka "outliers"

Many many properties of our real-world networks fit the above plot "shape"

→ power-law fit at the head

→ fat tail

Eg. component size distributions

degree distributions, "cluster"

↑ dense subgraph

size distribution

→ degrees vs. # of vertices w/ degree

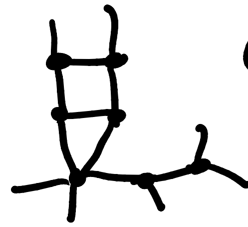
→ degrees vs. # of vertices w/ degree

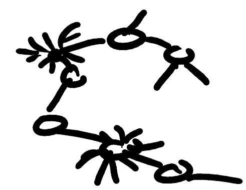
More on degree distributions

Why are they useful?

→ can capture alot of properties of the given network

E.g.: "skew" → gini coefficient
→ power law fit

Road network:  (not skewed)

social network:  (more skewed)

- Random graph generation and hypothesis testing

↳ Given some G with some degree distribution, how are the measured properties different than a random graph with the same degree distribution

---

More detail on power-law degree distributions

$-\gamma$ ← power-law exponent

$$P(k) \sim k^{-\gamma}$$ ← power-law exponent

proportional

probability of degree k

→ all together: frequency of a degree decreases exponentially as the degree increases

To actually compare degree distributions
(recall: measurements are relatively useless in a vacuum) (sp?)

- Calculate Gini coefficient
  ↳ straightforward
- Determine γ power-law exponent
  ↳ not as straightforward

One way to get γ: MLE

$$\gamma = 1 + n\left[\sum_{v \in V} \log\left(\frac{d(v)}{d_{min}}\right)\right]^{-1} \qquad \frac{n}{[\Sigma \dots]}$$

γ = power-law exponent

$d(v)$ = degree of v

$d(v)$ = degree of v

d_{min} = minimum over $d(v) \to 1$

$n = |V(G)|$ = number of verts in G

Most real-world graphs

$$1 \leq \gamma \leq 3$$

less skewed   more skewed

As we've noted:

this relation holds for more than simply degree distributions

Also: - connectivity size distributions
- cluster sizes
- shortest path lengths

All these together

=> <u>scale-free</u> graph

Generally, as we've stated:

[illegible] distributions with

Generally, as we've ...

→ Fat-tailed distribution with a power-law fit to the head of the distribution

Mining connectivity structure

How???

- Identifying cut vertices/edges
- k-(edge-) connectivity decomps for vary k
- look at component size distributions for varying k
- General reachability or how maximal k-components are connected

are connected

- $\underbrace{\text{k-core}}$ identification

  → maximal subgraph S where
  $\forall v \in V(S): d_S(v) \geq k$

  ↑ degree only considering edges to other $u \in V(S)$