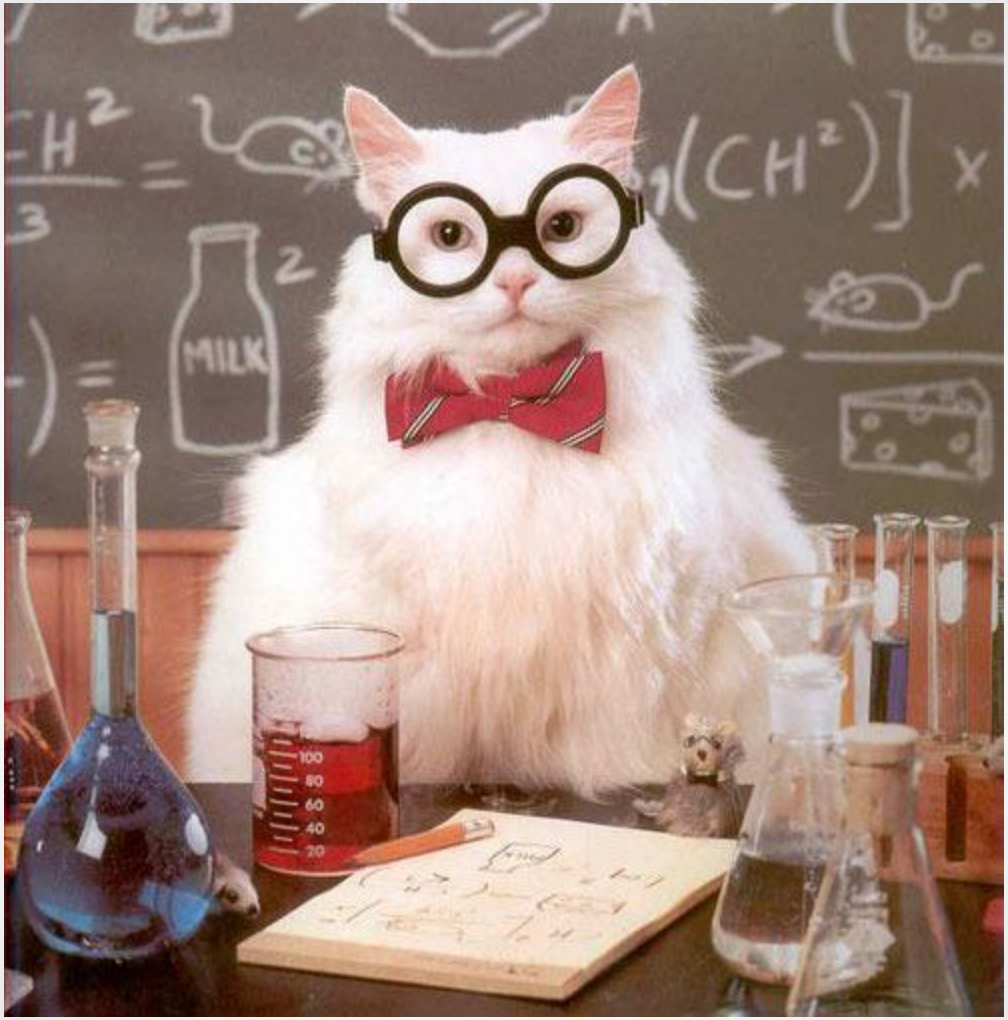
When the Exponent Matters

Marwan Burelle - LSE Summer Week 2015

Do you think P-Time algorithms



are tractable ?

LSE Security System
Laboratory of Epita

Numbers ...

	10	50	100	300	1000	10^6
5n	50	250	500	1500	5000	5×10^6
n × log n	33	282	665	2469	9966	14×10^6
n^2	100	2500	10000	90000	10^6	10^{12}
n^3	1000	125000	10^6	27×10^6	10^9	10^{18}
2^n	1024	$> 10^{15}$	$> 10^{30}$	$> 10^{90}$	$> 10^{301}$	too much

LSE Security System
Laboratory of Epita

10^{12} steps $\rightarrow$ 10 days

10^{18} steps $\rightarrow$ 300 centuries

LSE Security System
Laboratory of Epita

300 centuries ?

That’s long !

LSE Security System
Laboratory of Epita

Graphs

Used almost everywhere
Natural model for *networks* problems
Real graphs are big !

Graph Diameter

One out of many graph metrics
Linked to many other properties

LSE Security System
Laboratory of Epita

Diameter

- N: number of vertices
- M: number of edges $N \leq M \leq N^2$
- Real life sparse graphs: $M \sim N^{1+c}$
- Longest shortest path
- Naive algorithm: Warshall runs in $O(N^3)$
- BFS on adjacency lists:

  BFS: $O(N + M)$

  Diameter: $O(N^2 + N.M) = \Omega(N^2)$

Real Life Graph

- More than 10^6 vertices
- Sparse but connected

  $M = N^{1+c}$ with $0 \leq c < 1$
- No specific topology

You mean that diameter

takes days to compute ?

LSE Security System
Laboratory of Epita

Are we doomed ?

LSE Security System
Laboratory of Epita

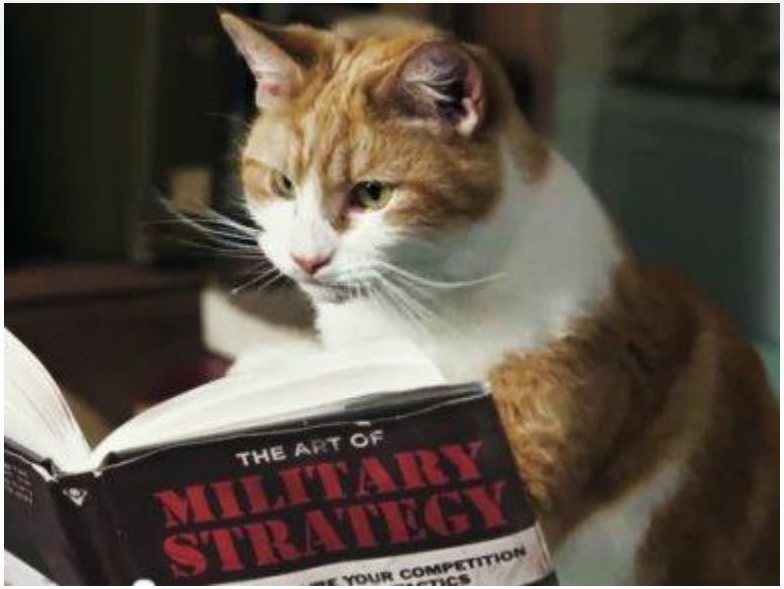
We can play with bounds

For any vertex v

$$\text{eccentricity}(v) \leq d \leq 2 \times \text{eccentricity}(v)$$

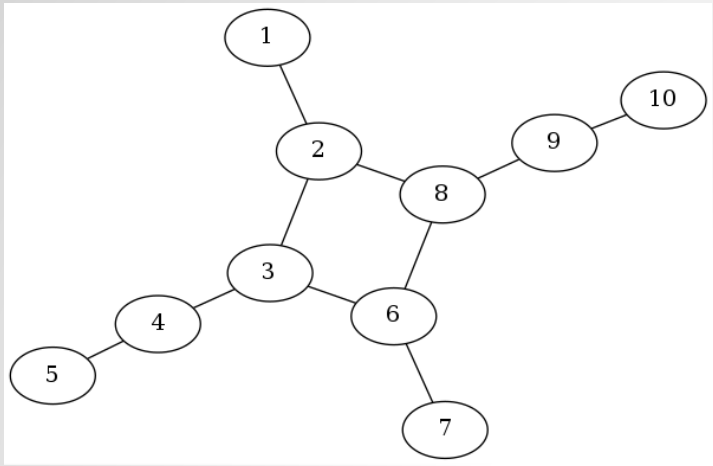
Still not enough:

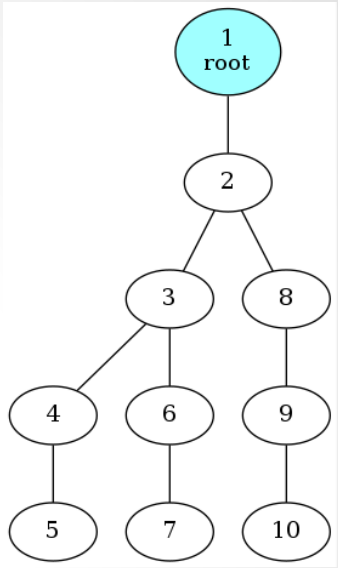
- can take times to collapse bounds
- may not converge
  - What if d is odd ?
  - Sometimes d < eccentricity(v)



Strategies

LSE Security System
Laboratory of Epita



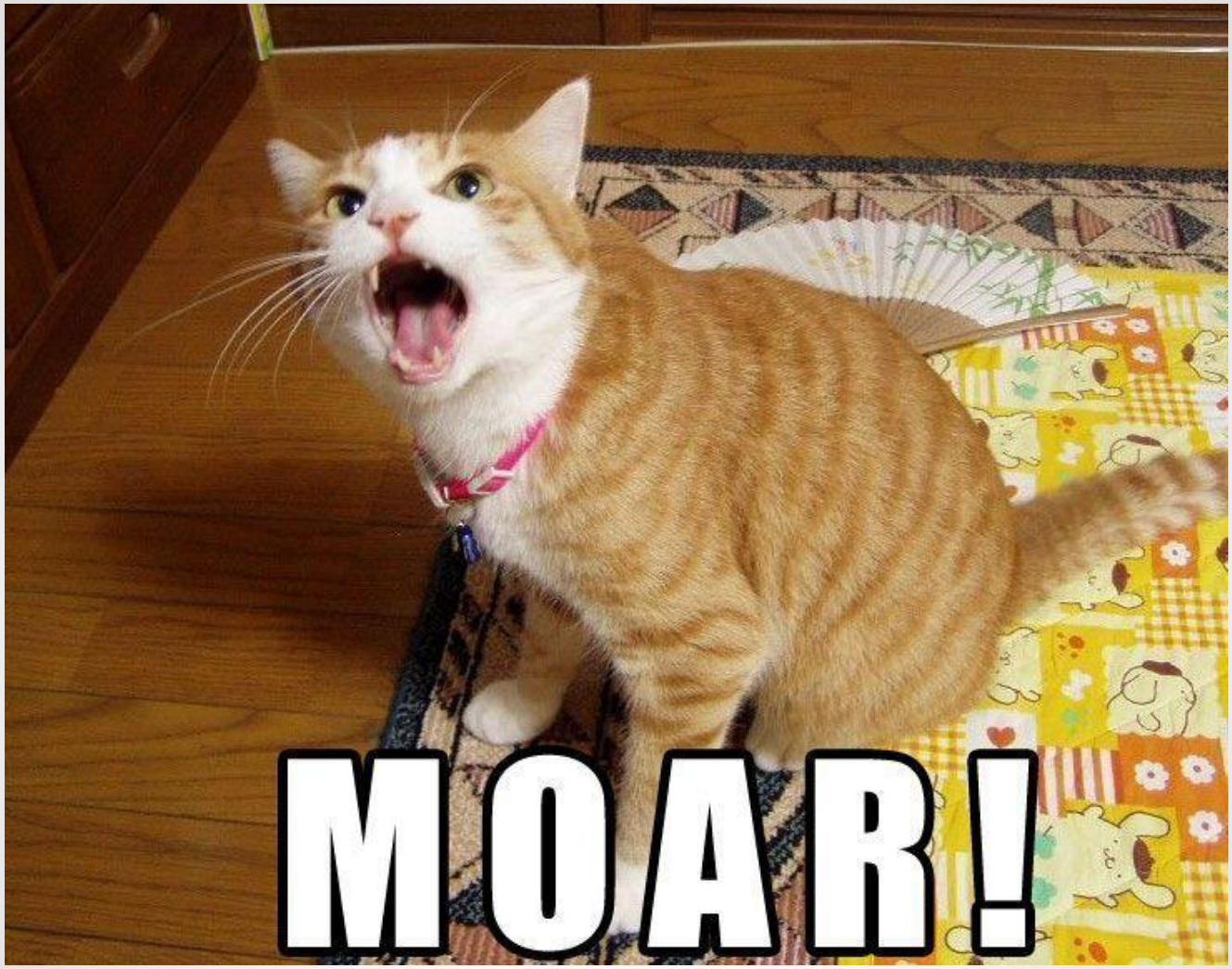


- BFS leaves contains diametral vertices
- Use intersection of leaves set

LSE Security System
Laboratory of Epita

Efficient for some cases

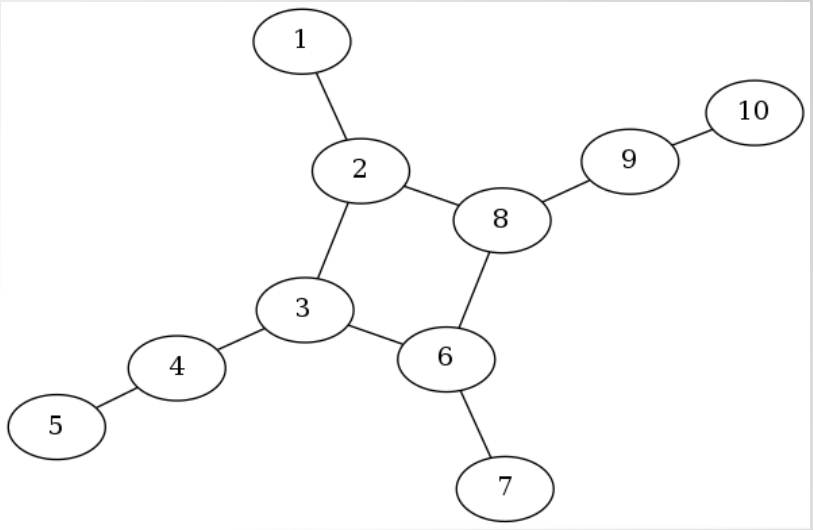
Sometimes leaves set is very stable

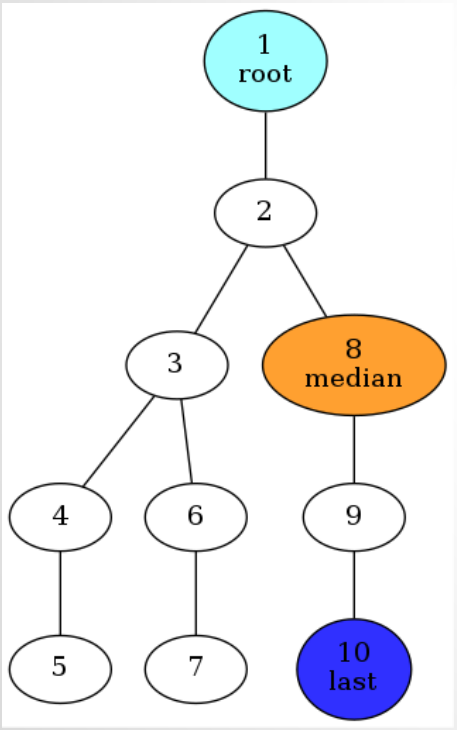


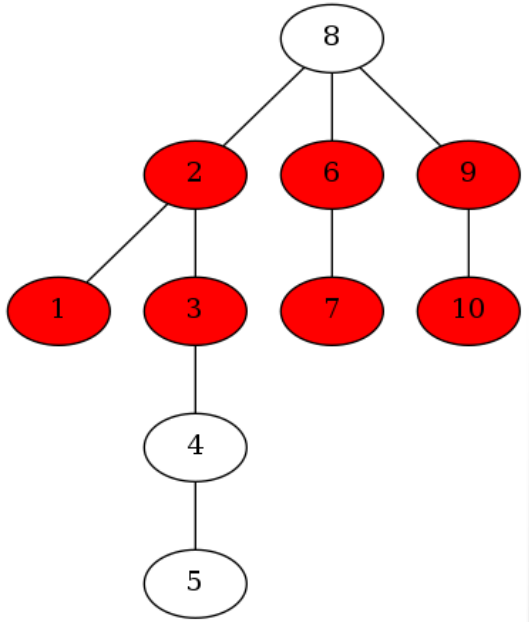
LSE
Security System
Laboratory of Epita

Eliminate more vertices:

- Use distance
- Use median point







LSE Security System
Laboratory of Epita

Initial vertex is important

- Use degree
- Use cut-vertices

Renumbering often helps

- Change encounter order
- Can improve memory access

Some results

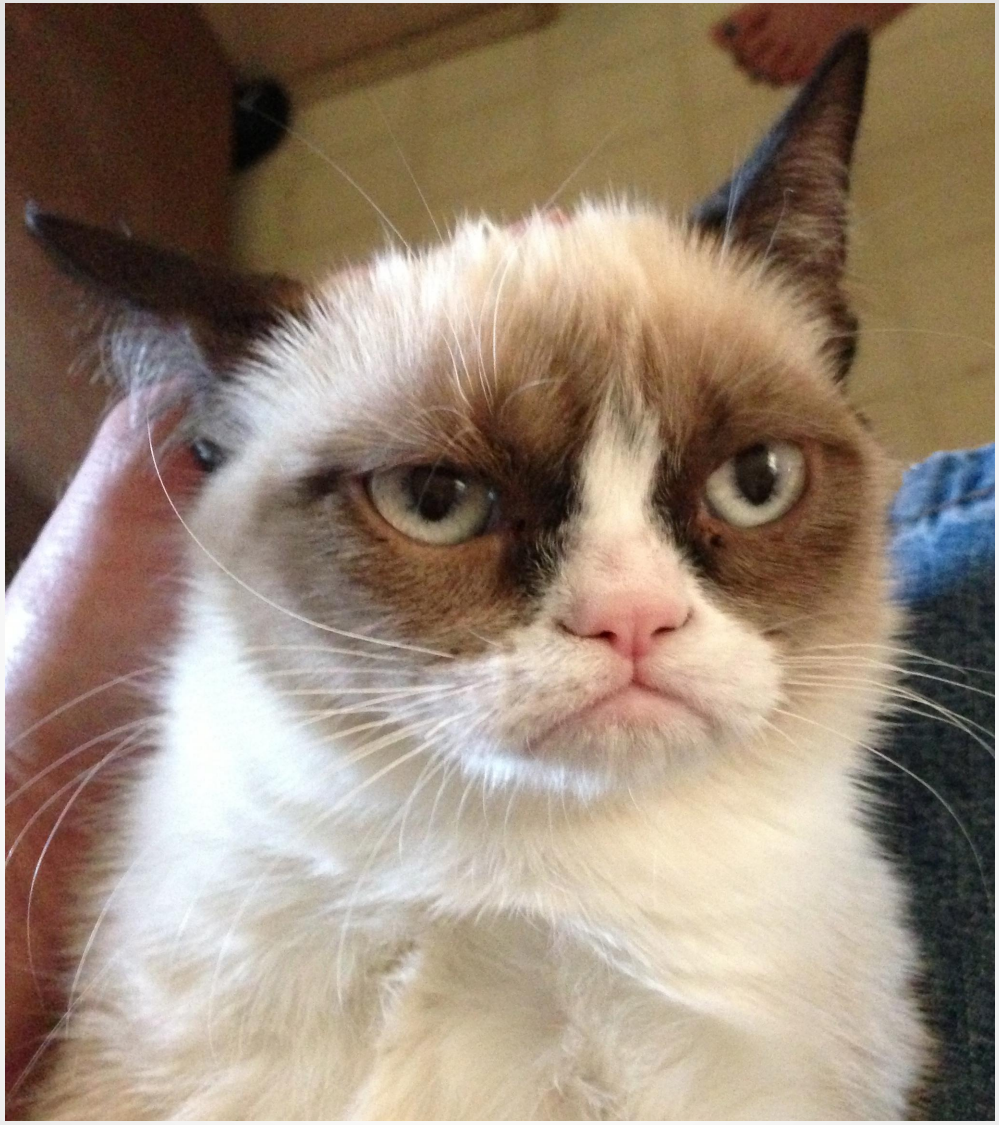
Graph	Order	Diameter	Runs	Lasagne
WEB	39459925	32	59	90.5
P2P	5792297	9	5	3588
roadNet-TX	1379917	1064	48	40246.30
finan512	74752	87	2129	29670.80

Lasagne: state of the art graph project
All tested graphs come from their page

http://piluc.dsi.unifi.it/lasagne/

More results published later, all but one are better with my code.

LSE Security System
Laboratory of Epita



Not bad ...

LSE Security System
Laboratory of Epita