Ultra-fast Rumor Spreading in Models of Real-World Networks

Konstantinos Panagiotou

(with N. Fountoulakis, T. Sauerwald)

MAX-PLANCK-GESELLSCHAFT

Information Dissemination

- Problem: Information Spreading in Networks
- Empirical observations:
  - Messages spread *extremely fast*, in particular when extraordinary events happen (natural disasters, plane crashes, …)
  - *Six Degrees of Separation*: Most vertices are close to each other (outliers exist) [Milgram ’60s, …]
- Question: How does the structure of real-world networks impact on the spread of information?

Rumor Spreading

- Initially a random vertex is informed
- Algorithm **SYNCH** (also known as PUSH-PULL)
  - Do in rounds:
  - Each *informed* vertex chooses a random neighbor, and sends the message to it
  - Each *uninformed* vertex chooses a random neighbor, and gets the information from it (if applicable)

[Karp et. al. ’00]

- Algorithm **ASYNCH**
  - Each vertex has an exponential clock with parameter 1
  - Whenever the clock of v rings:
  - If v is *informed*: transmit the information to a random neighbor
  - *Otherwise*: get the information from a random neighbor (if applicable)

Rumor Spreading (II)

- Typical question:
  - How long does it take *with high probability* until *everybody* is informed?
- Examples (SYNCH):
  - If the graph is complete with *n* vertices: $\sim \log_3 n$
  - For many graph classes: $\Theta(\log n)$
  - [Karp et. al. ’00, Panconesi et. al. ’10, Giakkoupis ‘11, Elsässer ’06, …]
  - Social Networks: later

A Different Question

- Observation from empirical studies:
  - In real networks there are *outliers*: these vertices have a large distance from almost all other vertices
  - *Effective diameter*: the minimum number of „hops“ in which 90% of vertices can reach each other [Faloutsos et al. ‘10]
- Question: How long does it take *whp* until 90% of the vertices are informed?

An ***1-ε*** fraction is informed?

„Bad“ News

- We say that a graph has *discrepancy D* if for every edge $\{u, v\}$

$$D^{-1} \leq \frac{\deg(u)}{\deg(v)} \leq D.$$

- **Lemma**. Whp, the number of informed vertices after *t* rounds is at most $2(2D+1)^t$.
- So, $\Omega(\log n)$ rounds are needed to inform $\sqrt{n}$ vertices if $D = O(1)$
- **Corollary**. „Many“ rounds are needed for complete graphs, random graphs, regular graphs, graphs with bounded maximum degree, ...

Models for Real-World Networks

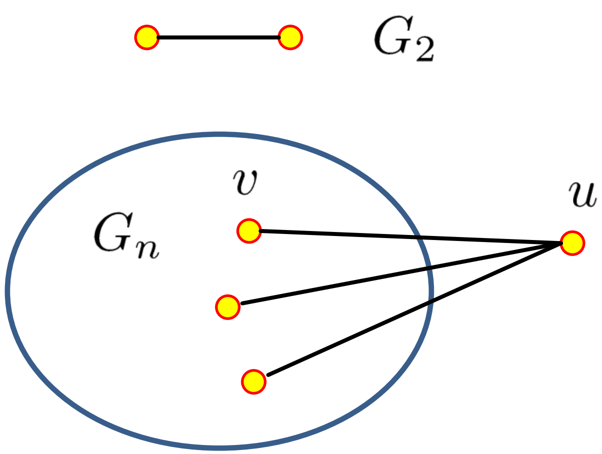
- Real World Networks do not look like classical random graphs
  - degree sequence is a *power law*:

$$c_1 k^{-\beta} n \leq N_k \leq c_2 k^{-\beta} n$$

  - Internet: $\beta \approx 2.2$, Facebook $\beta \approx 2.3$, WWW $\beta \approx 2.45$, …
  - Typical for social and technological networks: $2 < \beta < 3$
  - Non-vanishing *clustering coefficient*: many triangles
  - …

Models

- **Preferential Attachment (PA)**:
  Two parameters: *m*, $\delta > -m$



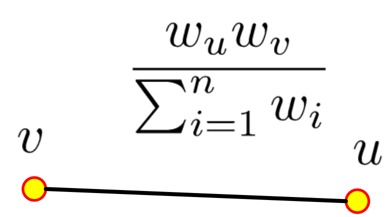
$\forall 1 \leq j \leq m$:

$\Pr[j\text{th edge of } u \text{ connects to } v] = \frac{deg(v)+\delta}{2e(G_n)+\delta n}$

$$N_k \sim c\, k^{-\beta}\, n, \quad \beta = 3 + \frac{\delta}{m}$$

[Barabasi, Albert '99, …]

- **Chung-Lu Random Graphs (CL)** (inhomogeneous random graphs):
- every vertex has a weight w_v
- Each edge is included independently with probability



- The expected degree of a vertex is $\sim w_v$

[Chung, Lu '03,
Bollobas, Janson, Riordan '07, …]

PA: The case $\delta = 0$

- Doerr, Fouz, Friedrich showed recently the following results for the PA model.
- **Theorem**. SYNCH spreads the rumor to all vertices in time $\Theta(\log n)$ with high probability.
- **Theorem.** There is a protocol CLEVER-SYNCH that spreads the rumor in time $\Theta(\frac{\log n}{\log \log n})$.

- Observation: spreading the rumor to $\sqrt{n}$ vertices requires also time $\Theta(\frac{\log n}{\log \log n})$.

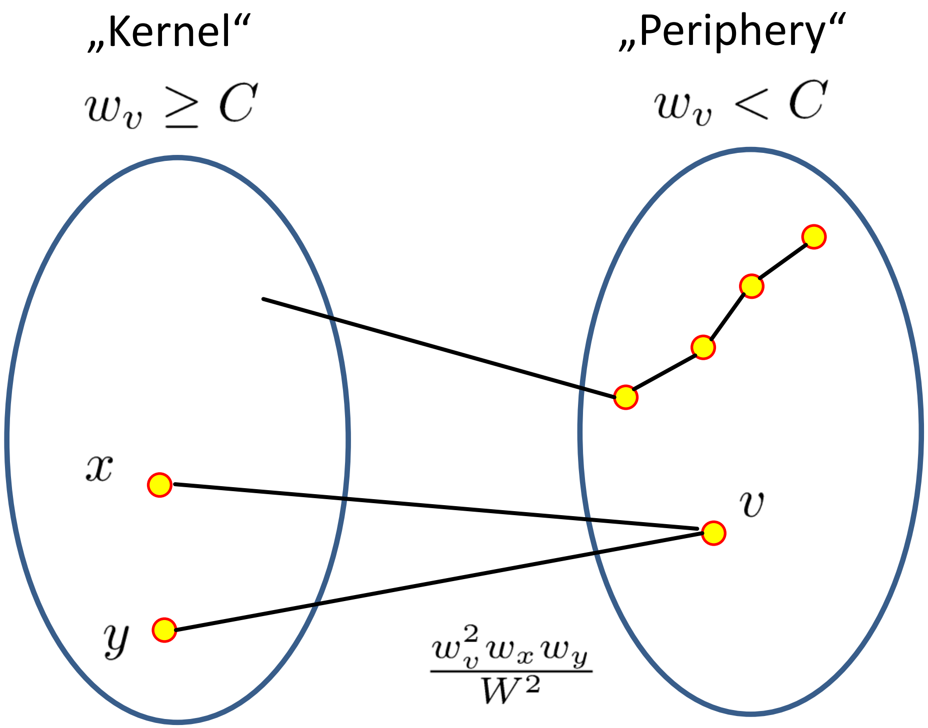
Our Results about SYNCH

- **Theorem** [Fountoulakis, P., Sauerwald]. Let *G* be either a PA graph or a CL graph with parameters such that the degree distribution is power law with exponent β. Then, with high probability
  - If $2 < \beta < 3$, then after $\Theta(\log \log n)$ rounds the rumor has spread to $n - o(n)$ vertices.
  - If $\beta > 3$, then $\Omega(\log n)$ are needed to spread the rumor to $\sqrt{n}$ vertices.

ASYNCH

- **Theorem**. [Fountoulakis, P., Sauerwald]. Let *G* be either a PA or a CL graph with $2 < \beta < 3$. Then, with high probability, after *a constant amount of time*, the roumor will spread to $(1-\epsilon)n$ vertices.
- **Conjecture**. Slower for $\beta \geq 3$.

Proof Idea



$$\approx \sum_{v: w_v < C} \frac{w_v^2 w_x w_y}{W^2} \approx \frac{w_x w_y}{n}$$

Theorem [Chung,Lu ‘03]. Almost all vertices are at distance *loglog(n)* from each other.

Efficient connector: v has degree 2

Auxiliary graph: contains a edge *xy* if there is an efficient connector

Claim. Remove every edge with probability ¼. Then 2*diameter of the resulting graph is an upper bound for the number of rounds.

Conclusions

- We showed that $2 < \beta < 3$ is special: rumor spreading needs *O(loglog(n))* rounds.
- All other (classical) models are *exponentially* slower.

- *Challenge*: perform a tight analysis on models that look more like real world networks.