



# Mixes on unstructured topologies

---

Shishir Nagaraja  
Computer Lab  
University of Cambridge

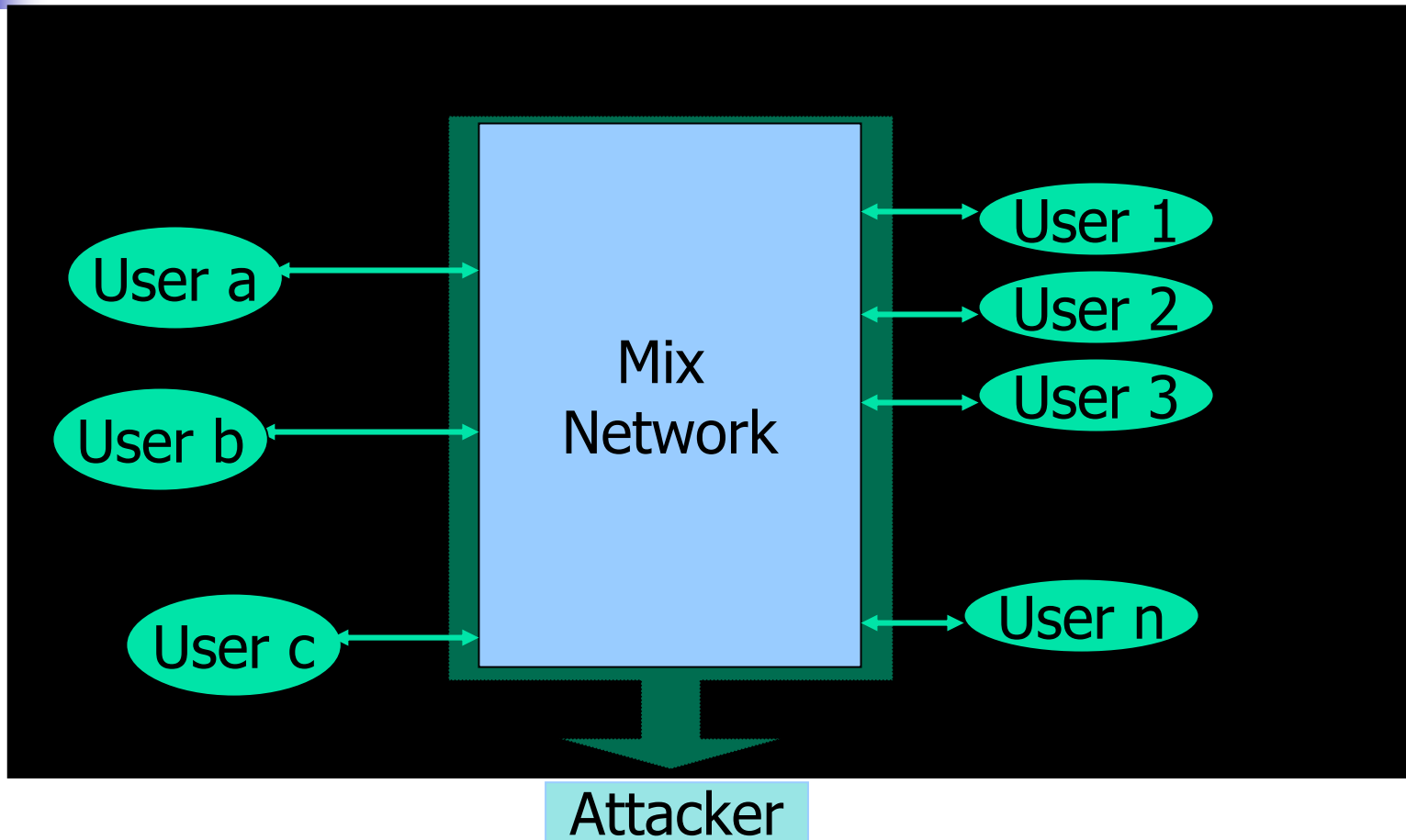


# what does anonymity mean?

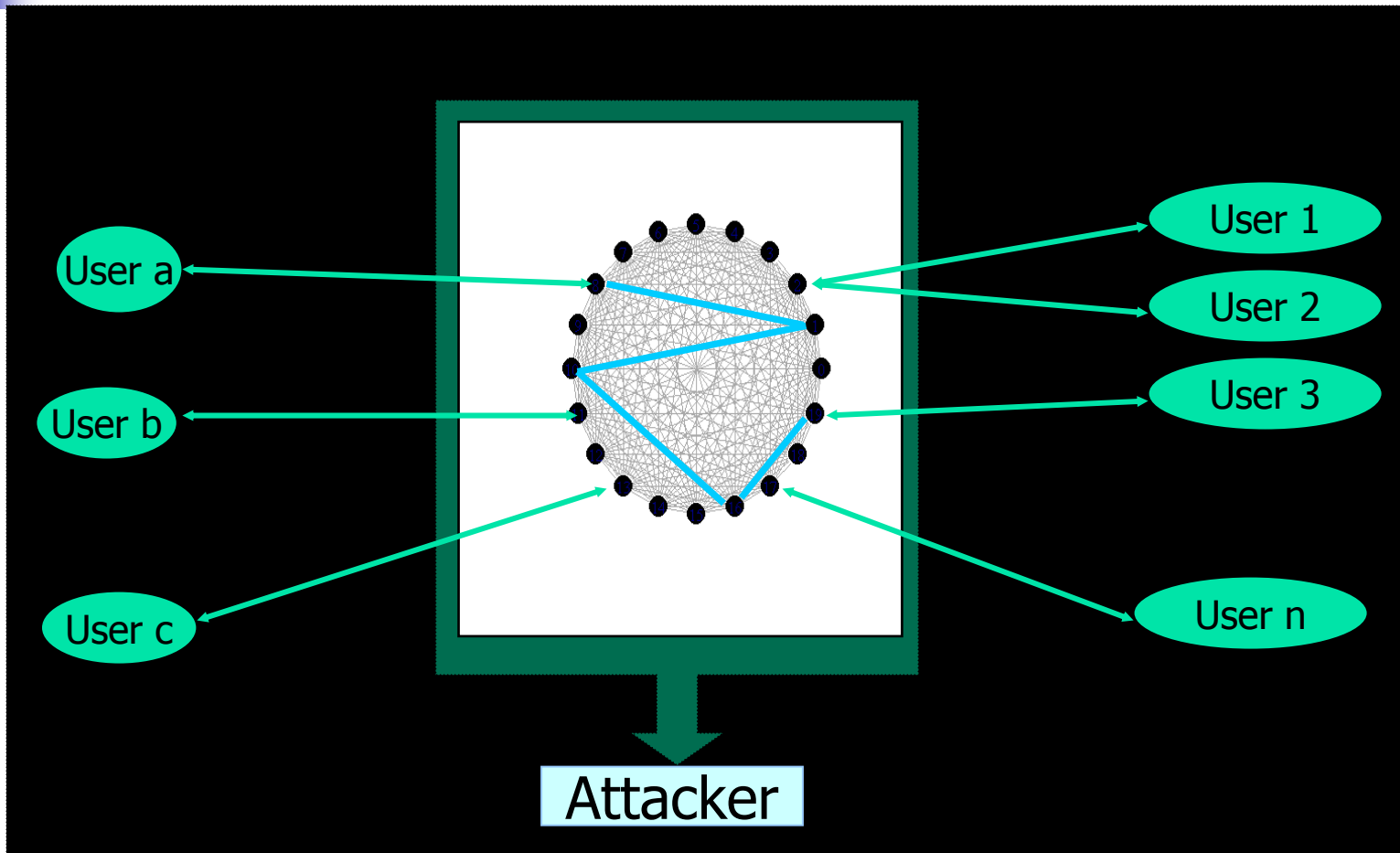
---

- **Unlinkability:** Hide the connection between the senders and the recipients.
- **Untraceability:** Hide the connection between actions of the same sender.
- **Unobservability:** Hide the fact that the user is talking.

# Mix-network



# Mix-network topology and mix-route



# Desirable properties of the mix-network



---

- 'High' traffic analysis resistance.
- Size: The larger the network, greater the anonymity set or maximal anonymity.
- Traffic: More traffic means better anonymity
- Robustness:
  - Liability management in anonymous communication
  - Clear incentives for carrying traffic under legal pressure



# Evaluation framework

---

- Is the given topology any good?
  - Figure out the efficiency of the mixing process.
  - Analyze the traffic-analysis resistance of the mix-network.
- Modeling mix network operation
  - Markovian random walks
- What we are not interested in:
  - Side channel analysis
  - Variation in protocol behaviour across topologies



# Measuring anonymity

---

- Number of bits the attacker is missing to uniquely link an actor to an action – (Serjantov and Danezis, PET 2002).

$$\mathcal{A} = \mathcal{E}[\alpha_i] = - \sum_i Pr[\alpha_i] \log_2 Pr[\alpha_i]$$



# Evaluation recap

---

Under conditions of maximal anonymity:

- Minimum mix-route length required.
- Amount of traffic needed to prevent intersection attacks – traffic load patterns.
- Resistance to corrupt nodes.

# Theory

1.

Walk of length  $t$

Infinite length Walk

$$\Delta(t) = \max_i \frac{|q_i^t - \pi_i|}{\pi_i}$$

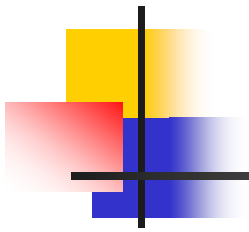
Convergence

$$\Delta(t) \leq n^{1.5} (\lambda_2)^t$$

Second eigenvalue of the transition matrix

$$T_{ij} = (1/k_i) A_{ij}$$

**A** Adjacency matrix  
**k<sub>i</sub>** Degree of node  $i$



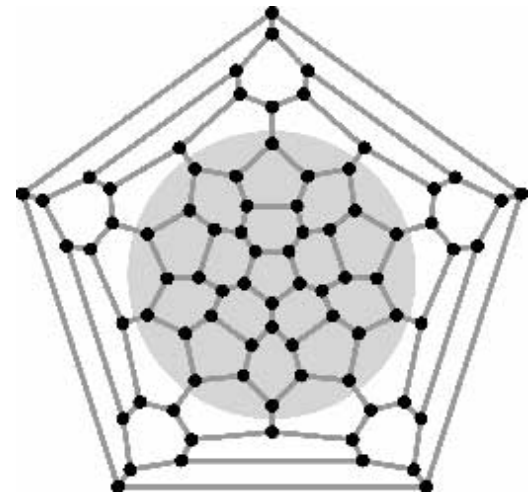
---

$$2. \quad q^t = q^{t-1}T$$

See the paper for a proof of why the second eigenvalue is a constant for varying network size ( $n$ ).

# Structured graph topologies

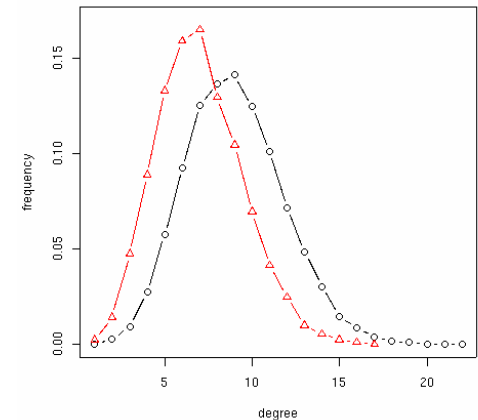
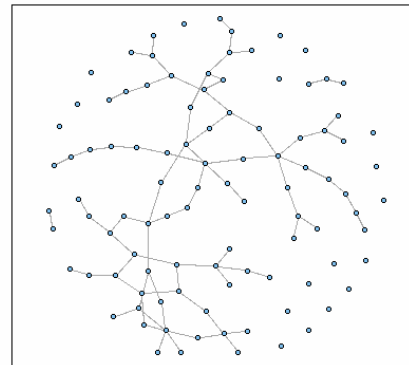
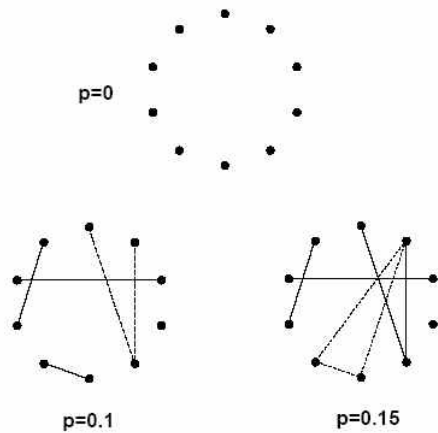
- Optimal mixing properties are obtained in expander graphs such as Ramanujan graphs.
- For  $N=5000$  nodes, we have,  $\lambda_2 \geq 0.5527$
- Hence, we can calculate mix-route length as approximately 4 hops.



(source:www.ams.org)

# Unstructured network topologies

- Erdős-Rényi random graph topology



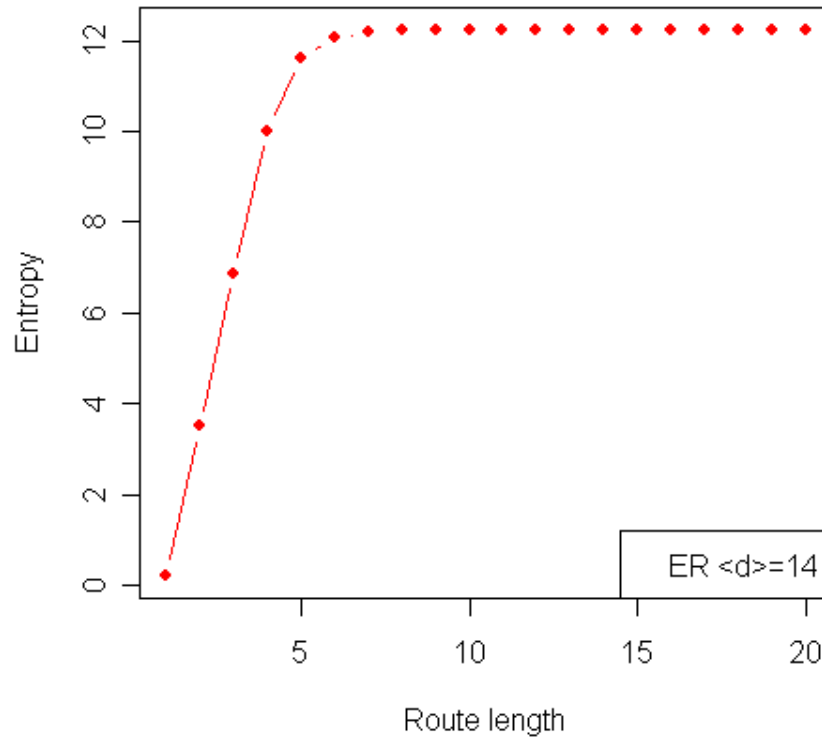
$$P(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

$$\lambda = \binom{N-1}{k} p_{ER}^k (1 - p_{ER})^{N-1-k}$$

We chose  $p$  such that the |biggest component|  $\sim |V|=5000$  nodes

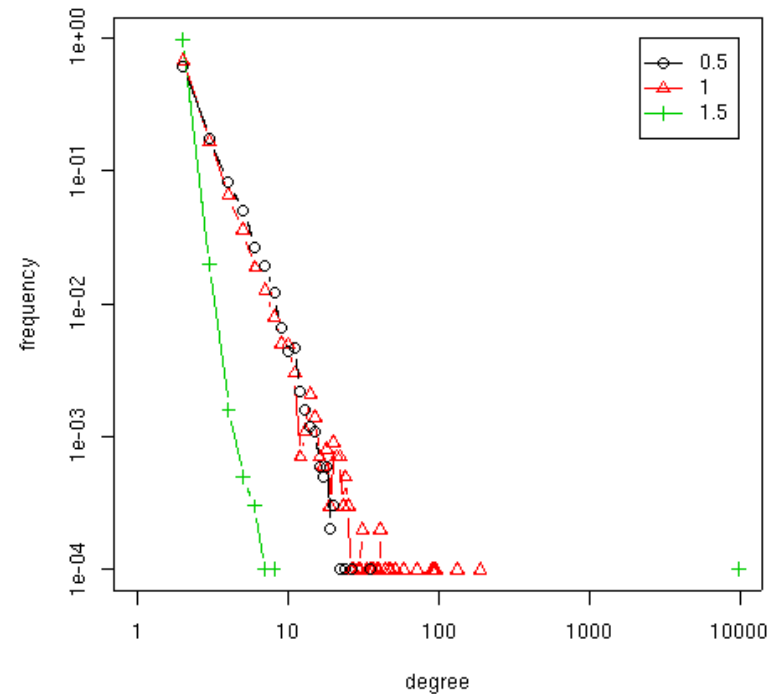
# Mixing efficiency of ER graph topology

**Convergence of random walks on ER networks**



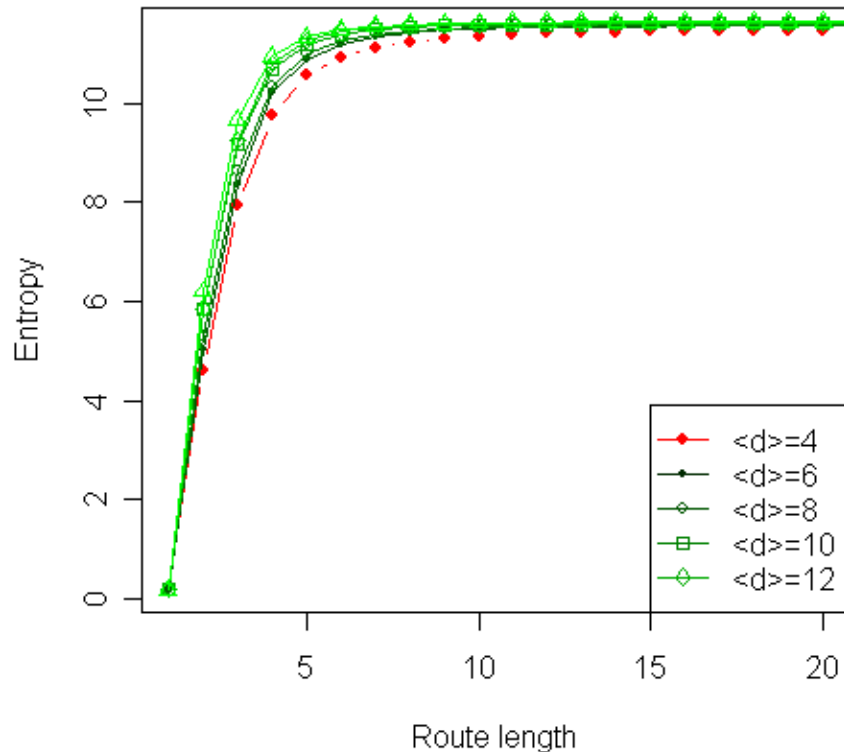
# Graph topologies continued...

- Scale-free topologies
- Power law
  - Heavy-tailed distribution
    - $P(X > x) \sim x^{-a}$ ,  $0 < a < 2$
  - Zipf distribution / Zeta distribution
    - $P(k) = Ck^{-(a+1)}$
  - Pareto distribution
    - $f(x) = ab^ax^{-(a+1)}$



# Mixing efficiency of SFR graph topology

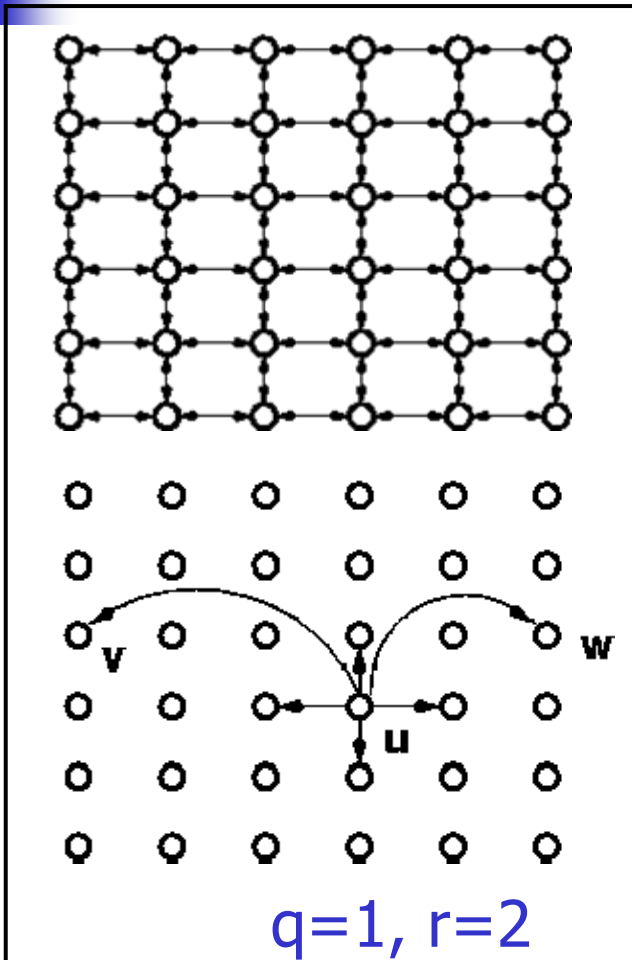
Convergence of random walks on scale-free random networks



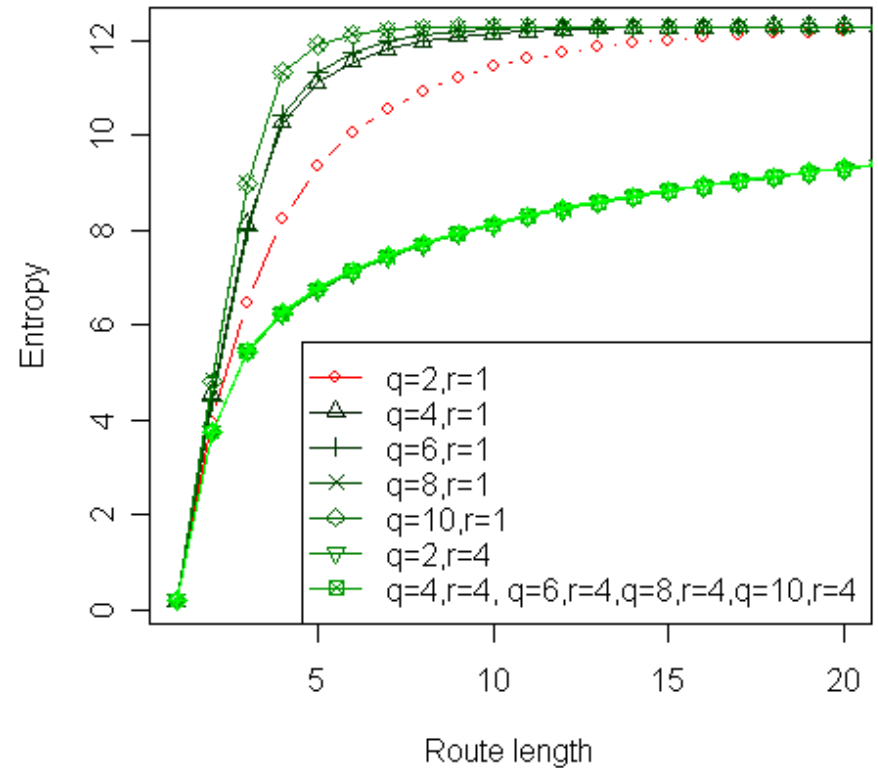
SFR models the massive AT&T call graph

Network ( $N = 5000$ )	$\langle d \rangle$ or $D$	$t$	$\mathcal{A}_{network}$
SFR	2	8	11.4383
	3	7	11.5626
	4	6	11.5958
	5	6	11.6135
	6	5	11.6351
ER	14	7	12.2339
Expander	14	4	12.2877

# Mixing efficiency of KWS topology (weak and strong ties)

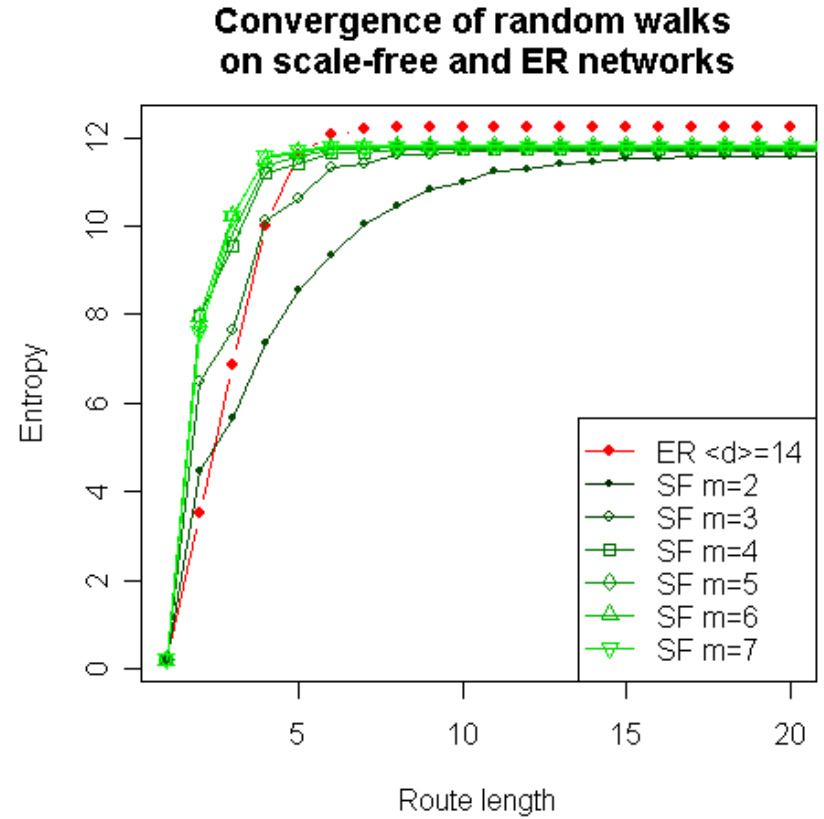
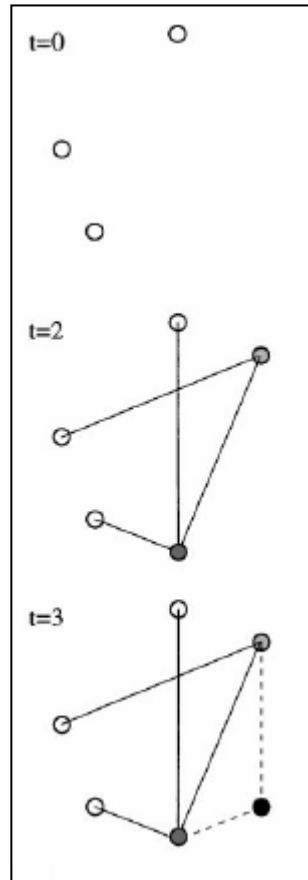


Convergence of random walks  
on Klienberg-Watts-Strogatz model



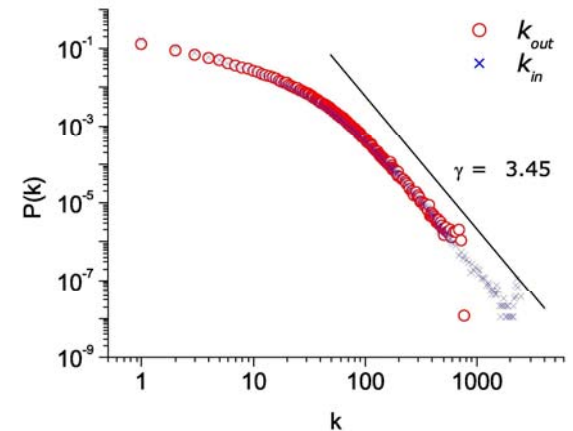
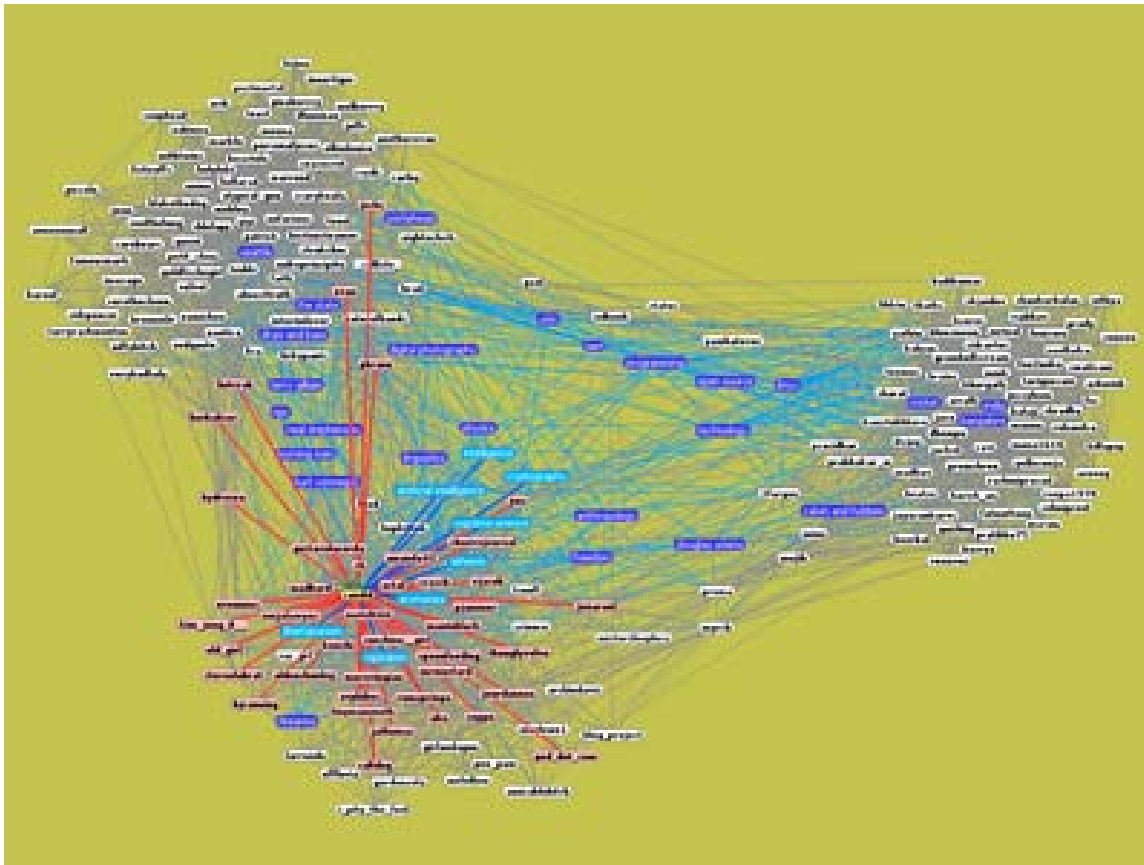
# Mixing efficiency of SF-BA graph topology

- Growth
  - Start with  $m_0$  nodes, and then add a node with  $m$  edges at every time step.
  - $m=m_0$
- Preferential attachment
 
$$\prod(k_i) = \frac{k_i}{\sum_j k_j}$$
  - It is a simple model but...
  - Fixed exponent = 3



# LiveJournal

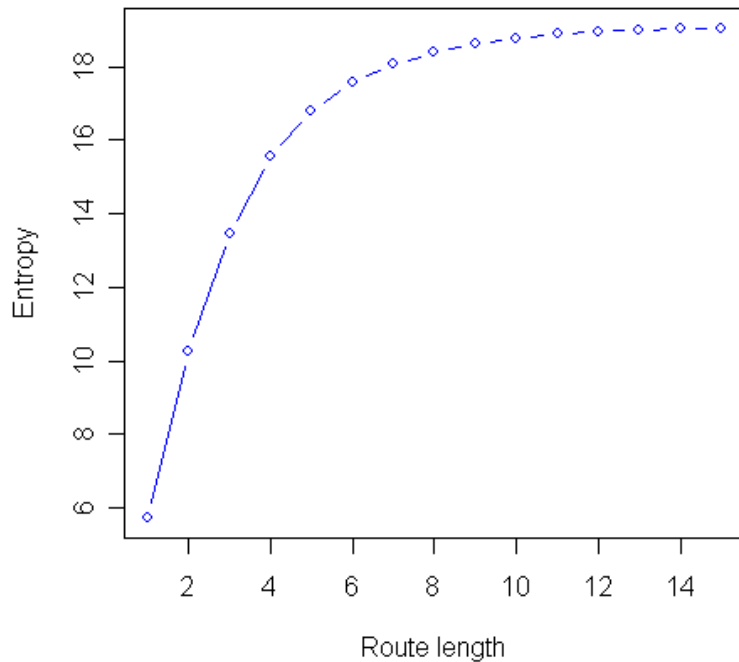
Source: Trejkaz Xaoza, Touchgraph



Pavel Zakharov, Thermodynamic approach for community discovering within the complex networks: LiveJournal study. e-print on arxiv.org: [physics/0602063](http://arxiv.org/abs/physics/0602063).

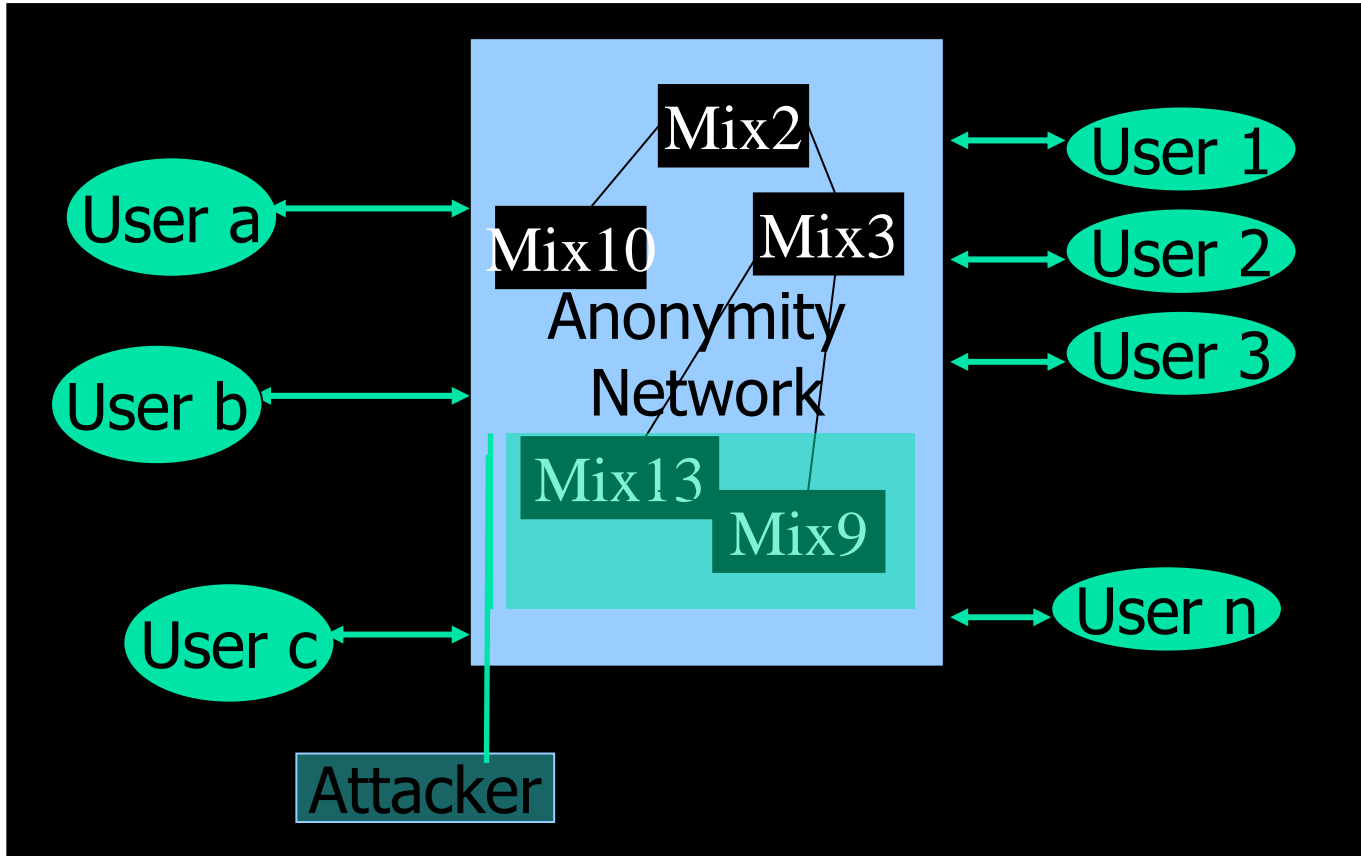
# Mixing efficiency of LiveJournal topology

Convergence of random walks on LiveJournal [N=3800000]



Network	$m$	$t$	$\mathcal{A}_{network}$
SF	2	15	11.5852
	3	10	11.6961
	4	6	11.7293
	5	6	11.7687
	6	6	11.7953
	7	6	11.8090
KWS	$q = 2, p = 1$	11	12.2945
	$q = 10, p = 1$	5	12.2939
	$q = 2, p = 4$	63	11.6440
	$q = 10, p = 4$	63	11.6380
ER	$\langle d \rangle = 14$	7	12.2339
Expander	$D = 14$	4	12.2877

# Corrupt nodes



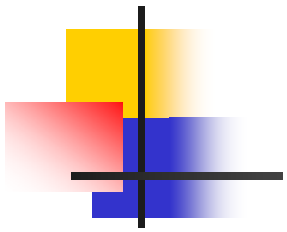


# Intersection attacks

---

- $\Pr[\text{any unused link}] \sim 0$
- Attacker: is the traffic from Alice proceeding along  $(i,j)$ ?
  - Mean volume vs observed volume of traffic
  - $L$  (confidence parameter) – number of standard deviations from the mean.
  - $b$  – batch size,  $p_i = 1/d_i$ ,  $k = 2$  mixing rounds.

$$k > 4L^2 \frac{p_i}{(1-p_i)} (b - 1)$$



Network	$\langle d \rangle$	$p_{min}$	Batch size	
SFR	2	0.0344	10.08	
	3	0.0222	15.84	
	4	0.0243	14.4	
	5	0.0192	18.36	
	6	0.0135	26.28	
	7	0.0125	28.44	■
KWS	27 ( $q = 1, r = 1$ )	0.0294	11.88	
	43 ( $q = 10, r = 1$ )	0.0169	20.88	
	26 ( $q = 1, r = 4$ )	0.0333	10.44	
	28 ( $q = 10, r = 4$ )	0.0294	11.88	
SF-linear	4	0.0048	74.16	
	6	0.0048	74.16	
	8	0.0041	86.04	
	10	0.0038	93.6	
	12	0.0037	96.12	
	14	0.0031	112.32	■
LJ	7.3221	0.00857	41.64	■
ER	14	0.0333	10.44	■
Expander	14	0.0714	4.68	■

Table 1: Batch sizes required to prevent intersection attacks



## structured vs unstructured...

---

- Traffic analysis resistance – **Comparable**
- Maximal anonymity – **Comparable**
- Topological robustness in the face of litigation pressure – **this depends on the social capital in the network.**
  - Friends process each other's traffic.
  - Processing "3<sup>rd</sup> party" traffic - indirect reciprocity - encourages a diverse user base which brings its benefits (Anonymity loves company, Dingledine and Mathewson 2006).



# Conclusions

---

- A successful mix-network design needs to consider the issues of liability management.
- Tapping social capital in a network to enhance topological robustness is an attractive proposal,
- And we have established the essential technical feasibility of this if this means using an unstructured mixnet topology
- Specifically:
  - Mix-route length is not a problem
  - Corrupt hubs are not a problem either
  - Batch sizes are however a challenge. Where do we bring 8 times the amount of dummy traffic from when compared to expanders?
    - will social chatter suffice?