Peer-to-Peer Networks
08 Kelips and Epidemic Algorithms

Christian Schindelhauer
Technical Faculty
Computer-Networks and Telematics
University of Freiburg

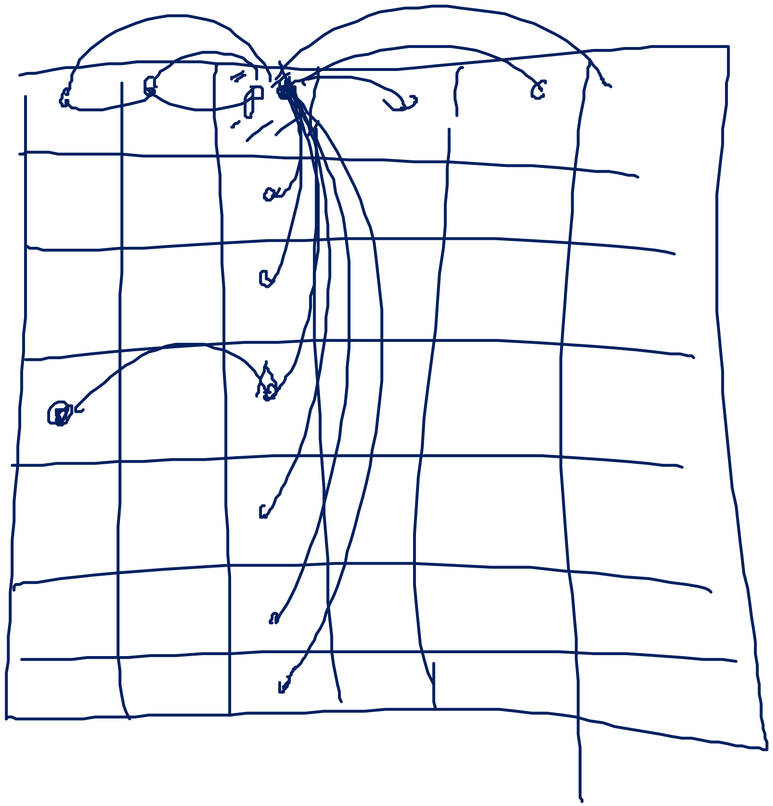
Kelips

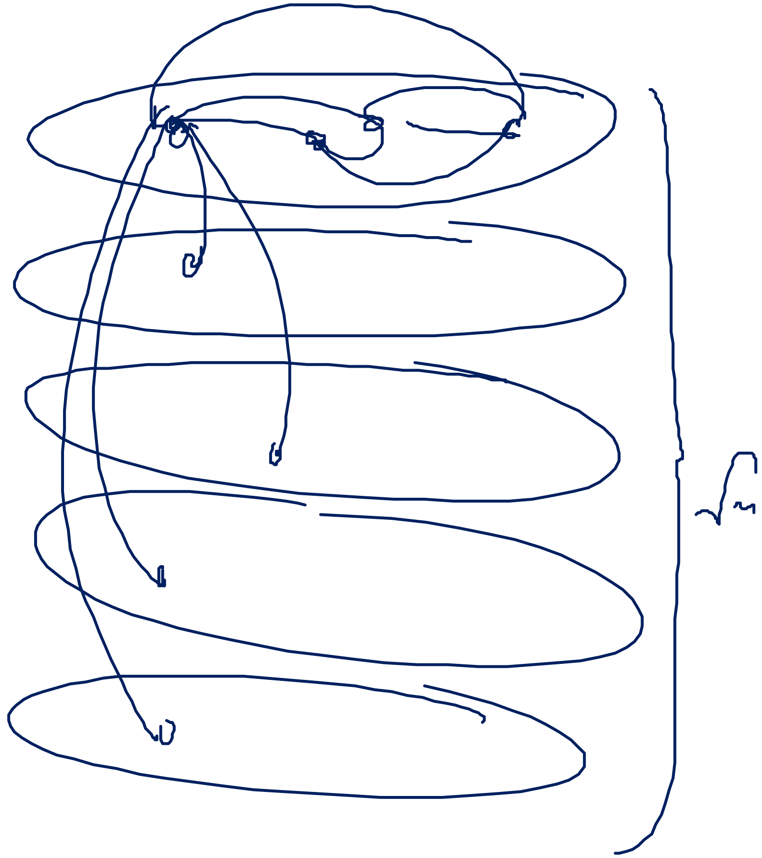
- Indranil Gupta, Ken Birman, Prakash Linga, Al Demers, Robbert van Renesse
  - Cornell University, Ithaca, New York
- Kelip-kelip
  - malay name for synchronizing fireflies
- P2P Network
  - uses DHT
  - constant lookup time
  - $O(n^{1/2})$ storage size
  - fast and robust update

	Lookup	out-degree
CAN	$O(n^{1/d})$	$\sim O(d)$
Chord	$O(\log n)$	$O(\log n)$
Gnutella	?	~ Pareto ?
Napster	$O(1)$	$n-1$
Koorde	$O(\log n)$	$O(1)$
Kelips	$O(1)$	$O(\sqrt{n})$

d: dimension

$2 \cdot m = 2 \cdot \sqrt{n}$ $\qquad$ $n = m^2$



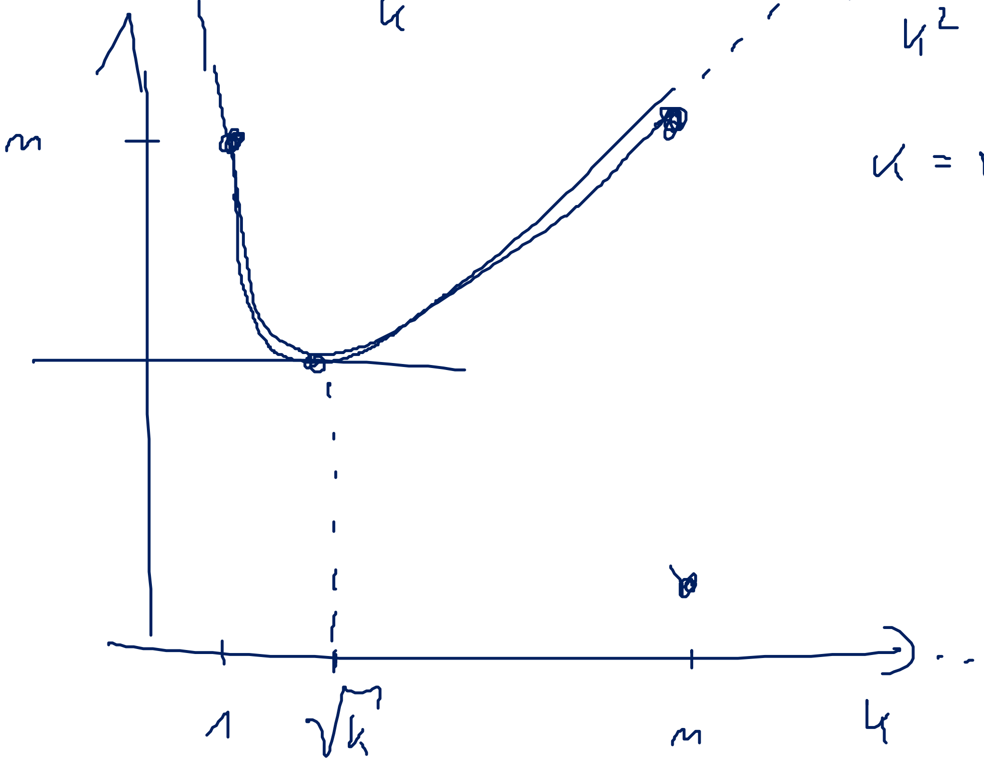


$$\frac{d\,f(k)}{dk} = 0$$

$$f(k) = \frac{n}{k} + k$$

$$-\frac{n}{k^2} + 1 = 0$$

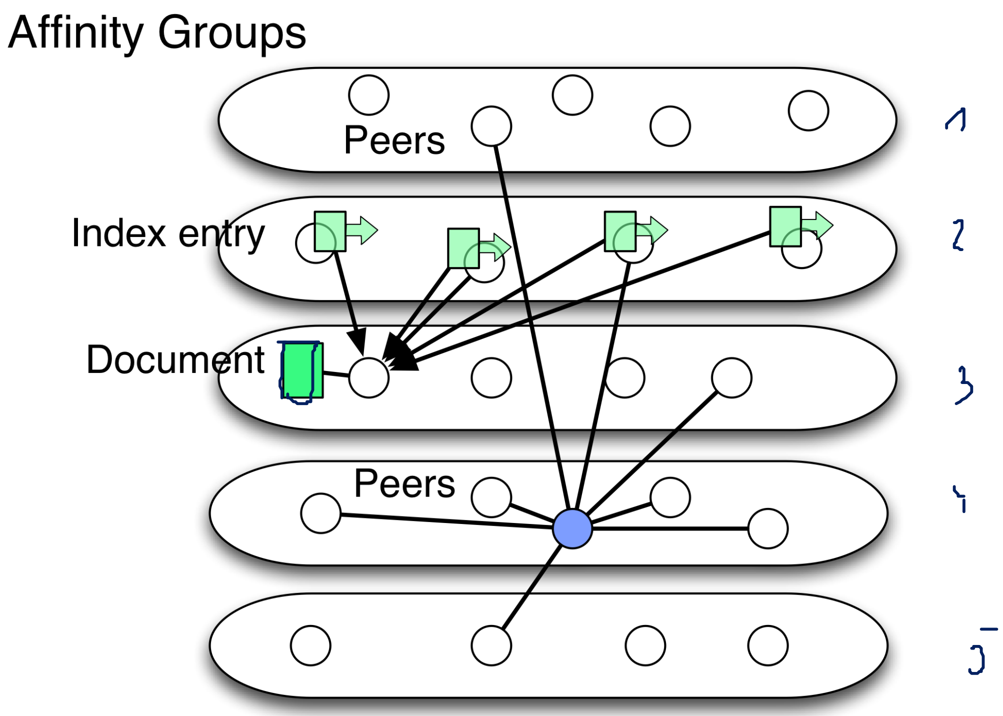
$$k = \sqrt{n}$$



Kelips Overview

$\sqrt{10^6} = \boxed{1000}$

- Peers are organized in k affinity groups
  - peer position chosen by DHT mechanism
  - k is chosen as $n^{1/2}$ for n peers
- Data is mapped to an affinity group using DHT
  - all members of an affinity group store all data
- Routing Table
  - each peer knows all members of the affinity group
  - each peer knows at least one member of each affinity group
- Updates
  - are performed by epidemic algorithms

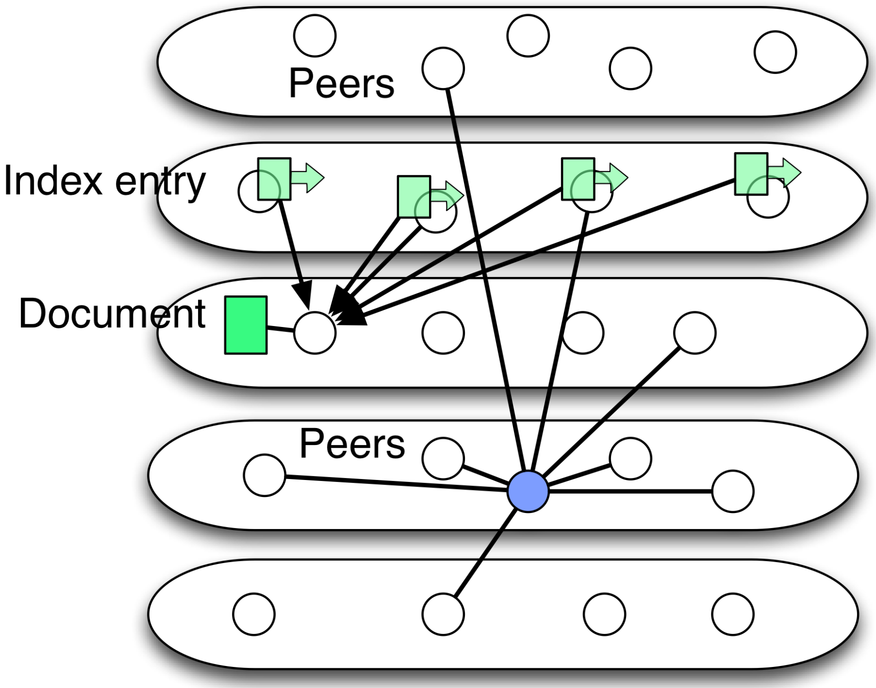


Routing Table

CoNe Freiburg

- Affinity Group View
  - Links to all O(n/k) group members
  - This set can be reduced to a partial set as long as the update mechanism works
- Contacts
  - For each of the other affinity group a small (constant-sized) set of nodes
  - O(k) links
- Filetuples
  - A (partial) set of tuples, each detailing a file name and host IP address of the node storing the file
  - O(F/k) entries, if F is the overall number of files
- Memory Usage: O(n/k + k + F/k)
  - for $k = O(\sqrt{n+F})$

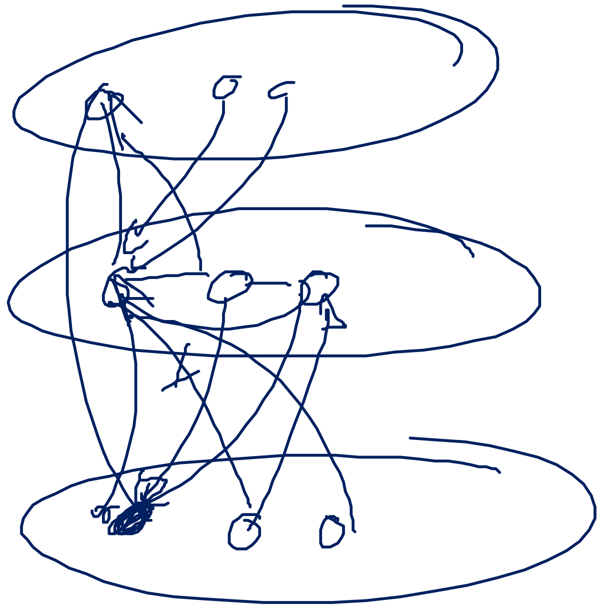
Affinity Groups



$O(\sqrt{n+F})$

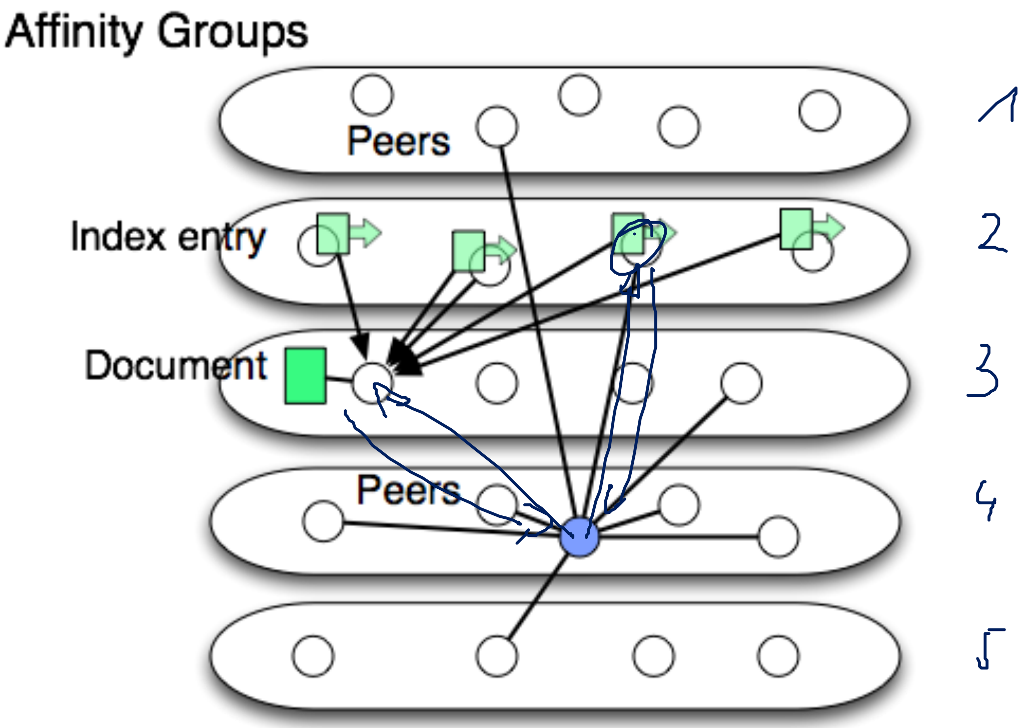
$\boxed{\frac{n}{k} + k} + \frac{F}{k}$

UNI FREIBURG



Lookup

- Lookup-Algorithm
  - compute index value
  - find affinity group using hash function
  - contact peer from affinity group
  - receive index entry for file (if it exists)
  - contact peer with the document
- Kelips needs four hops to retrieve a file

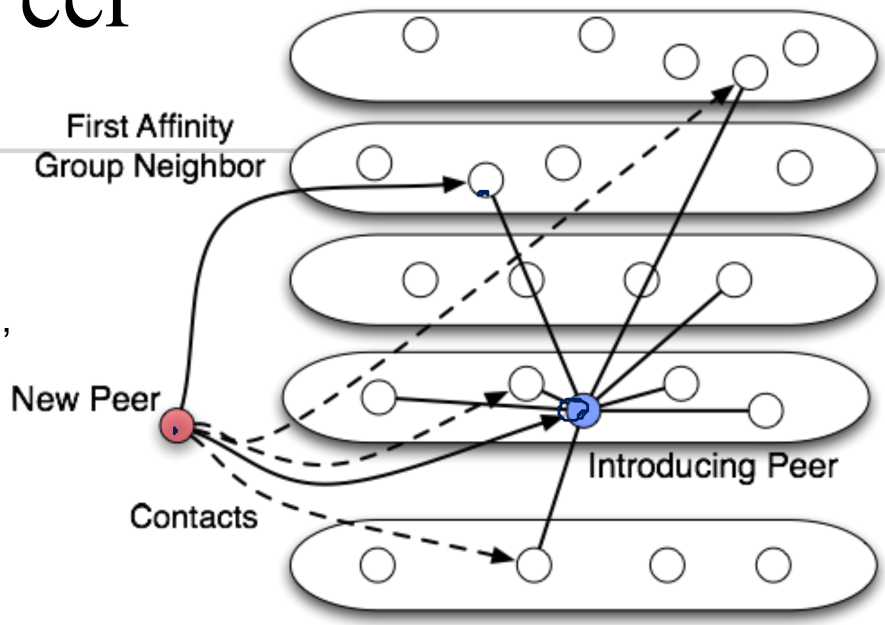


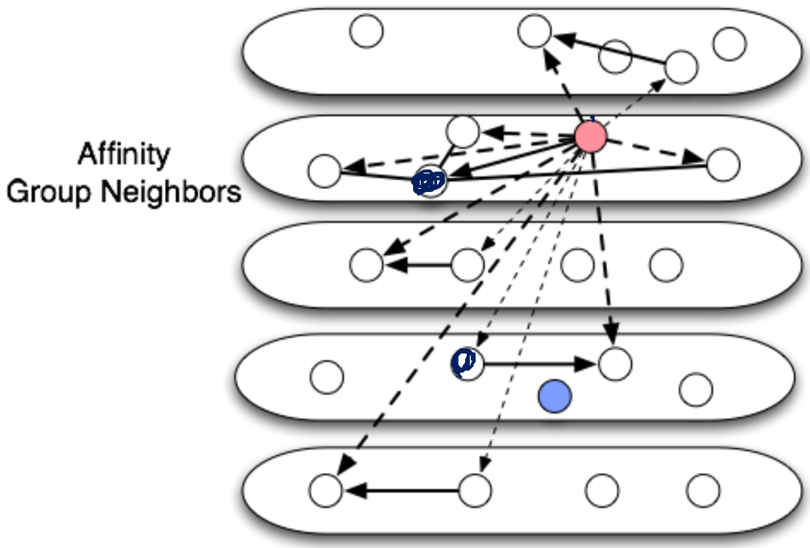
h("Lady Gaga") = 2

Inserting a Peer

CoNe Freiburg

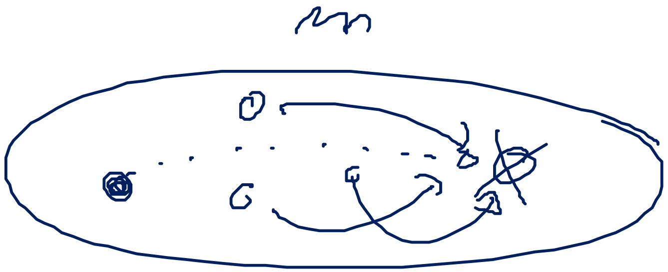
- **Algorithm**
  - Every new peer is introduced by a special peer, group or other method,
    - e.g. web-page, forum etc.
  - The new peer computes its affinity group and contacts any peer
  - The new peer asks for one contact of the affinity group and copies the contacts of the old affinity group
  - By contacting a neighbor node in the affinity group it receives all the necessary contacts and index filetuples
  - Every contact is replaced by a random replacement (suggested by the contact peer)
  - The peer starts an epidemic algorithm to update all links
- Except the epidemic algorithm the runtime is O(k) and only O(k) messages are exchanged





UNI FREIBURG

$O(\sqrt{n})$



$$\frac{1}{n} = \frac{1}{\sqrt{n}}$$

$$\frac{1}{4} \le \left(1 - \frac{1}{n}\right)^n \le \frac{1}{e} \approx \frac{1}{2.718\ldots}$$

$$\underbrace{\frac{1}{e} \cdot \frac{1}{e} \cdots\cdots \frac{1}{e}}_{r}$$

$$\left(\frac{1}{e}\right)^r \le \frac{1}{n^c} \Rightarrow r \le O(\log n)$$

$$r = c \ln n$$

How to Add a Document

- Start an Epidemic Algorithm to Spread the news in the affinity group
- Such an algorithm uses O(n/k) messages and needs O(log n) time
- We introduce Epidemic Algorithms later on

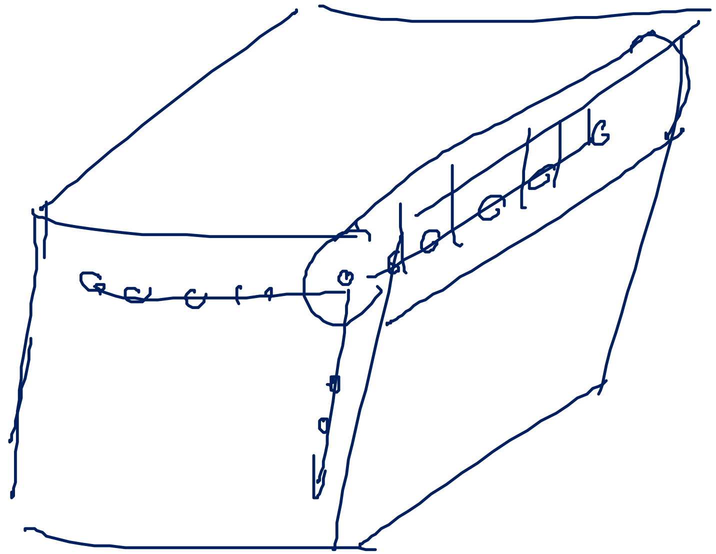
How to Check Errors

- Kelip works in heartbeats, i.e. discrete timing
- In every heartbeat each peer checks one neighbor
- If a neighbor does not answer for some time
  - it is declared to be dead
  - this information is spread by an epidemic algorithm
- Using the heartbeat mechanisms all nodes also refresh their neighbors
- Kelips quickly detects missing nodes and updates this information

Discussion

- Kelips has lookup time $O(1)$, but needs $O(n^{1/2})$ sized Routing Table
  - not counting the $O(F/n^{1/2})$ file tuples
- Chord, Pastry & Tapestry use lookup time $O(\log n)$ but only $O(\log n)$ memory units
- Kelips is a reasonable choice for medium sized networks
  - up to some million peers and some hundred thousands index entries

degree: $O(n^{1/3}) = O(\sqrt[3]{n})$ with constant hops



$3\sqrt[3]{n} - 1$

To Do

- What is an Epidemic Algorithm

Fibonacci

Epidemic Spread of Viruses

- Observation
  - a - most viruses do not prosper in real life
  - b - other viruses are very successful and spread fast
- How fast do viruses spread?
- How many individuals of the population are infected?
- Problem
  - a - social behavior and infection risk determine the spread
  - the reaction of a society to a virus changes the epidemy
  - viruses and individuals may change during the infection

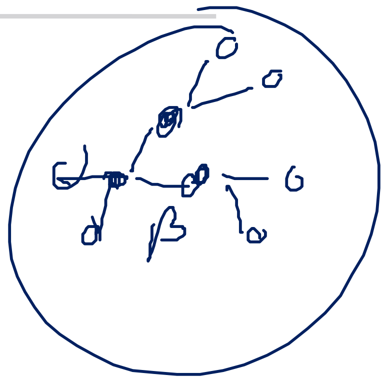
Mathematical Models

- SI-Model (rumor spreading)
  - susceptible → infected
- SIS-Model (birthrate/deathrate)
  - susceptible → infected → susceptible
- SIR-Model
  - susceptible → infected → recovered
- Continuous models
  - deterministic
  - or stochastic
- Lead to differential equations
- Discrete Models
  - graph based models
  - random call based
- Lead to the analysis of Markov Processes

CoNe Freiburg

Infection Models

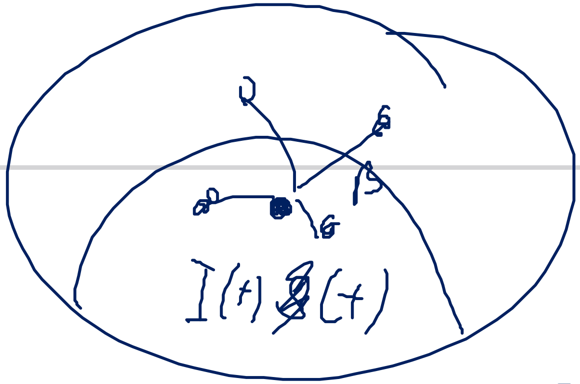
S → I



- SI-Model (rumor spreading)
  - susceptible → infected
  - At the beginning one individual is infected
  - Every contact infects another indiviual
  - In every time unit there are in the expectation ß contacts
- SIS-Model (birthrate/deathrate)
  - susceptible → infected → susceptible
  - similar as in the SI-Model, yet a share of δ of all infected individuals is healed and can receive the virus again
  - with probability δ an individual is susceptible again
- SIR-Model
  - susceptible → infected → recovered
  - like SI-Model, but healed individuals remain immune against the virus and do not transmit the virus again

CoNe Freiburg

SI-Model



$\frac{S(t)}{n} = s(t)$

$\beta \cdot s(t) \cdot I(t)$

- Variables
  - n: total number of individuals
    - remains constant
  - S(t): number of (healthy) susceptible individuals at time t
  - I(t): number of infected individuals
- Relative shares
  - s(t) := S(t)/n
  - i(t) := I(t)/n
- At every time unit each individual contacts ß partners
- Assumptions:
  - Among ß contact partnres ß s(t) are susceptible
  - All I(t) infected individuals infect ß s(t) I(t) other individuals in each round
- Leads to the following recursive equations:
  - I(t+1) = I(t) + ß s(t) I(t)
  - i(t+1) = i(t) + ß i(t) s(t)
  - S(t+1) = S(t) – ß s(t) I(t)
  - s(t+1) = s(t) – ß i(t) s(t)

UNI FREIBURG

CoNe Freiburg

SI-Model

$s(t) = 1 - i(t)$

- i(t+1) = i(t) + ß i(t) s(t) $= i(t) + \beta \cdot i(t) \cdot (1 - i(t))$
- s(t+1) = s(t) – ß i(t) s(t)
- Idea:
  - i(t) is a continuous function
  - i(t+1)-i(t) approximate first derivative

$i(t+1) - i(t) = \beta \cdot i(t) \cdot (1 - i(t))$

$$\frac{i(t+1)-i(t)}{1} \approx \frac{di(t)}{dt}$$

$$\boxed{\frac{di(t)}{dt} = \beta \cdot i(t)(1-i(t))}$$

- Solution:

$$i(t) = \frac{1}{1+\underbrace{\left(\frac{1}{i(0)}-1\right)}_{c}e^{-\beta t}}$$

$$\frac{1 + c \cdot e^{-\beta t} - c \cdot e^{-\beta t}}{1 + c \cdot e^{-\beta t}} = 1 - \frac{c \cdot e^{-\beta t}}{1 + c \cdot e^{-\beta t}}$$

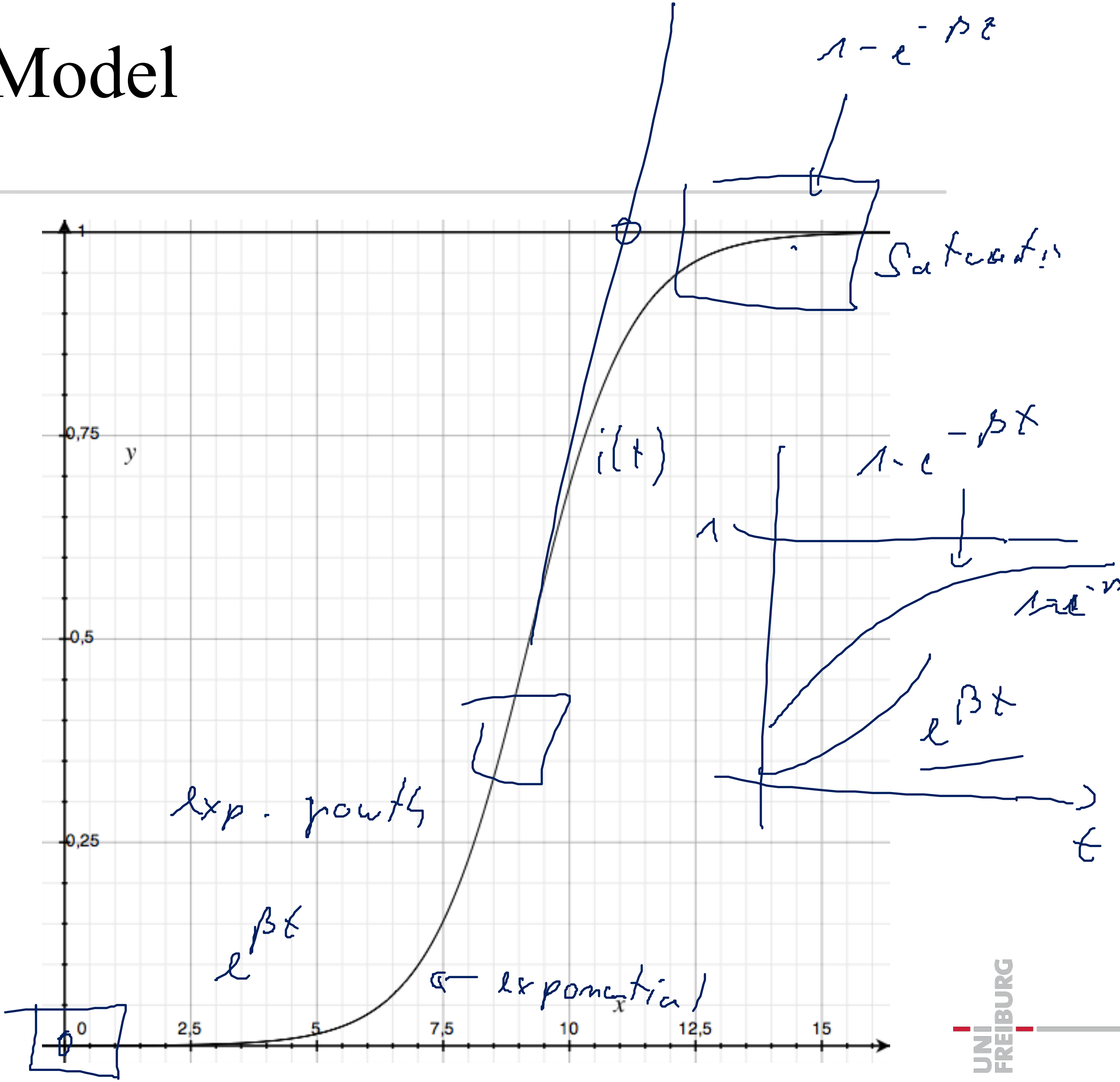
$$\frac{1}{e^{-\beta t}} = e^{\beta t} \approx 0$$

UNI FREIBURG

CoNe Freiburg

SI-Model

- The number of infected grows exponentially until half of all members are infected
- Then the number of susceptible decrease exponentially



UNI FREIBURG

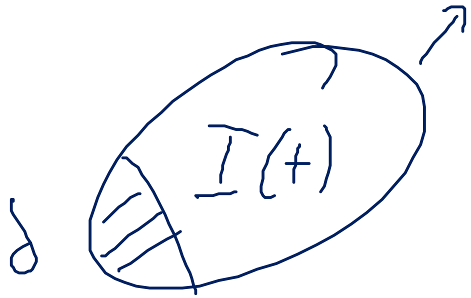
SIS-Model

- **Variables**
  - n: total number of individuals
    - remains constant
  - S(t): number of (healthy) susceptible individuals at time t
  - I(t): number of infected individuals
- **Relative shares**
  - s(t) := S(t)/n
  - i(t) := I(t)/n
- At every time unit each individual contacts ß partners

CoNe Freiburg

SIS-Model

- **Assumptions**:
  - Among ß contact partners ß s(t) are susceptible
  - All I(t) infected individuals infect ß s(t) I(t) other individuals in each round
  - A share of δ of all infected individuals is susceptible again
- Leads to the following recursive equations:
  - $I(t+1) = I(t) + ß\, i(t)\, S(t) - \delta\, I(t)$
  - $i(t+1) = i(t) + ß\, i(t)\, s(t) - \delta\, i(t)$
  - $S(t+1) = S(t) - ß\, i(t)\, S(t) + \delta\, I(t)$
  - $s(t+1) = s(t) - ß\, i(t)\, s(t) + \delta\, i(t)$



UNI FREIBURG

SI-Model

1 - i(t)

- i(t+1) = i(t) + ß i(t) s(t) – δ i(t)
- s(t+1) = s(t) – ß i(t) s(t) + δ i(t)
- Idea:
  - i(t) is a continuous function
  - i(t+1)-i(t) approximate first derivative $\frac{i(t+1)-i(t)}{1} \approx \frac{di(t)}{dt}$

$$\frac{di(t)}{dt} = \beta \cdot i(t)(1-i(t)) - \delta i(t)$$

1

- Solution:
  - for $\rho = \frac{\delta}{\beta}$

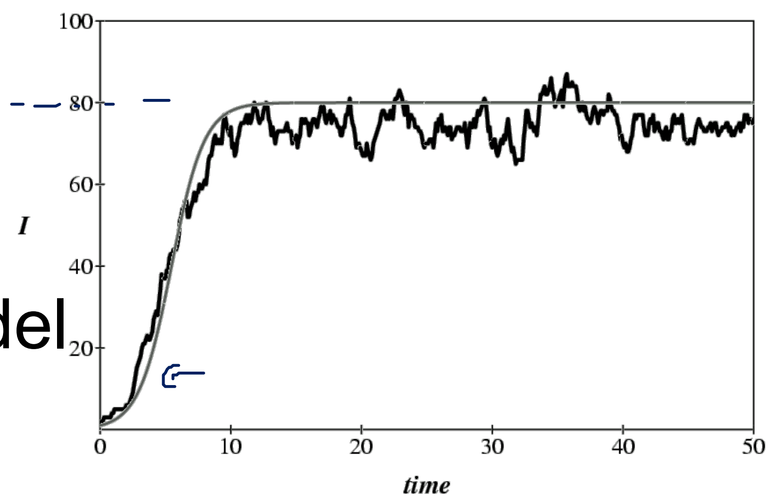
$$i(t) = \underbrace{\frac{1-\rho}{1+\left(\frac{1-\rho}{i(0)}-1\right)e^{-(\beta-\delta)t}}}_{C}$$

SIS-Model
Interpretation of Solution

$$i(t) = \frac{1-\rho}{1 + \left(\frac{1-\rho}{i(0)} - 1\right) e^{-(\beta-\delta)t}} \qquad \rho = \frac{\delta}{\beta}$$

- If ß < δ
  - then i(t) is strictly decreasing
- If ß > δ
  - then i(t) converges against $1 - \rho = 1 - \delta/ß$
- Same behavior in discrete model has been observed
  - [Kephart,White'94]

$1 - \rho$

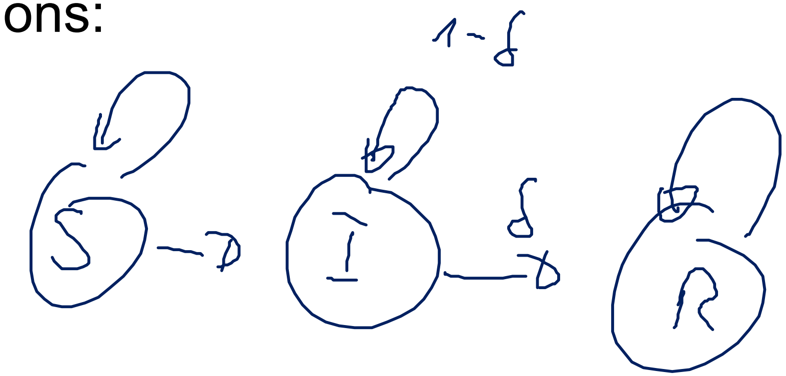


SIR-Model

- Variables
  - n: total number of individuals
    - remains constant
  - S(t): number of (healthy) susceptible individuals at time t
  - I(t): number of infected individuals
  - R(t): number or recovered individ.
- Relative shares
  - s(t) := S(t)/n
  - i(t) := I(t)/n
  - r(t) := R(t)/n
- At every time unit each individual contacts ß partners

SIR-Model

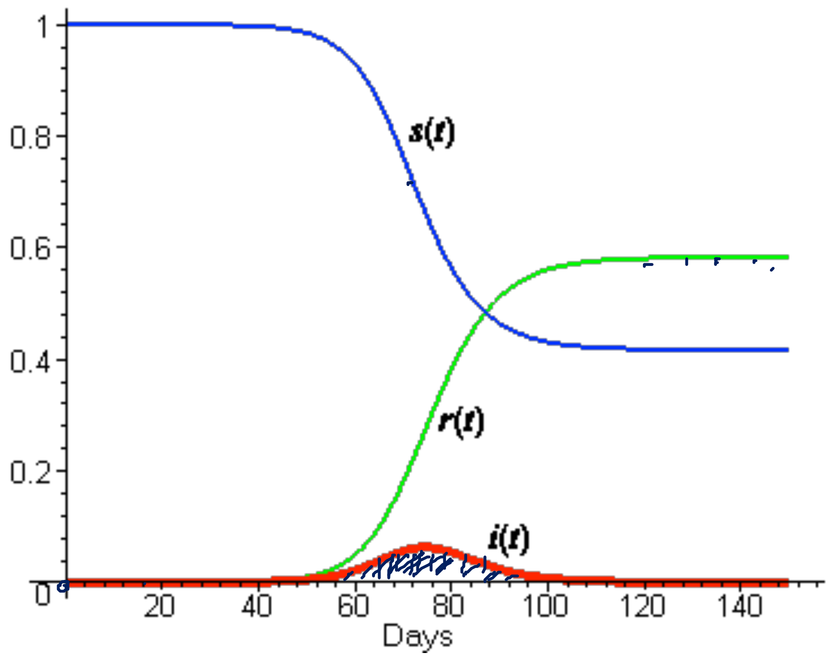
- Assumptions:
  - Among ß contact partners ß $s(t)$ are susceptible
  - All $I(t)$ infected individuals infect ß $s(t)\, I(t)$ other individuals in each round
  - A share of δ of all infected individuals is immune (recovered) and never infected again
- Leads to the following recursive equations:
  - $I(t+1) = I(t) + \text{ß}\, i(t)\, S(t) - \delta\, I(t)$
  - $i(t+1) = i(t) + \text{ß}\, i(t)\, i(t) - \delta\, i(t)$
  - $S(t+1) = S(t) - \text{ß}\, i(t)\, S(t)$
  - $s(t+1) = s(t) - \text{ß}\, i(t)\, s(t)$
  - $R(t+1) = R(t) + \delta\, I(t)$
  - $r(t+1) = r(t) + \delta\, i(t)$



SIR-Model

- The equations and its differential equations counterpart
  - $i(t+1) = i(t) + ß\, i(t)\, i(t) - \delta\, i(t)$
  - $s(t+1) = s(t) - ß\, i(t)\, s(t)$
  - $r(t+1) = r(t) + \delta\, i(t)$
- No closed solution known
  - hence numeric solution
- Example
  - $s(0) = 1$
  - $i(0) = 1.27 \cdot 10^{-6}$
  - $r(0) = 0$
  - $ß = 0.5$
  - $\delta = 0.3333$

$$\frac{ds(t)}{dt} = -\beta \cdot i(t)s(t)$$
$$\frac{di(t)}{dt} = \beta \cdot i(t)s(t) - \delta i(t)$$
$$\frac{dr(t)}{dt} = -\delta i(t)$$

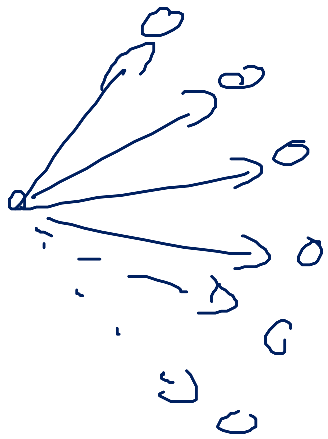


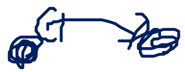
CoNe Freiburg

Replicated Databases

DNS

- Same data storage at all locations
  - new entries appear locally
- Data must be kept consistently
- Algorithm is supposed to be decentral and robust
  - since connections and hosts are unreliable
- Not all databases are known to all
- Solutions
  - Unicast
    - New information is sent to all data servers
  - Problem:
    - not all data servers are known and can be reached
  - Anti-Entropy
    - Every local data server contacts another one and exchanges all information
    - total consistency check of all data
  - Problem
    - comunication overhead
- Epicast …





UNI FREIBURG

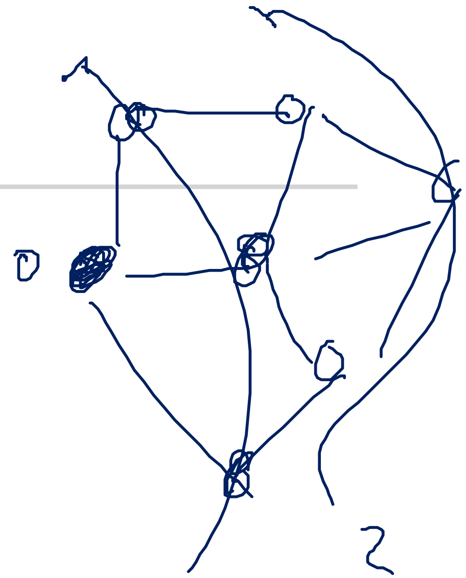
Epidemic Algorithms

Rumor Spreading

- Epicast
  - new information is a rumor
  - as long the rumor is new it is distributed
  - Is the rumor old, it is known to all servers
- Epidemic Algorithm [Demers et al 87]
  - distributes information like a virus
  - robust alternative to BFS or flooding
- Communication method
  - Push & Pull, d.h. infection after $\log_3 n + O(\log \log n)$ rounds with high probability
- Problem:
  - growing number of infections increases comunication effort
  - trade-off between robustness and communication overhead

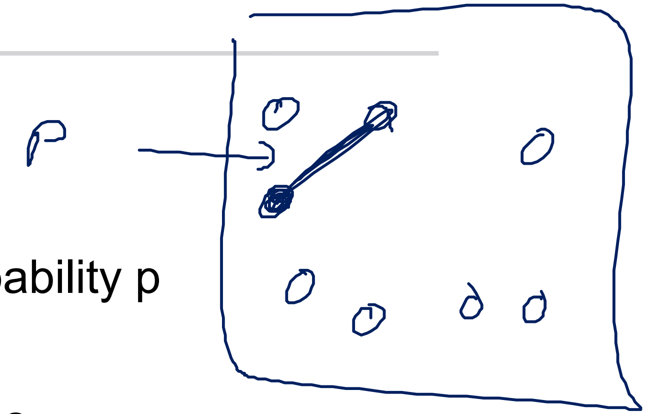
SI-Model for Graphs

- Given a contact graph G=(V,E)
  - n: number of nodes
  - I(t) := number of infected nodes in round t
  - i(t) = I(T)/n
  - S(t) := number of susceptible nodes in round t
    - I(t)+S(t)=n
  - s(t) = S(T)/n
- Infection:
  - If u is infected in round t and $(u,v) \in E$, then v is infected in round t+1
- Graph determines epidemics
- Complete graph:
  - 1 time unit until complete infection
- Line graph
  - n-1 time units until complete infection



BFS

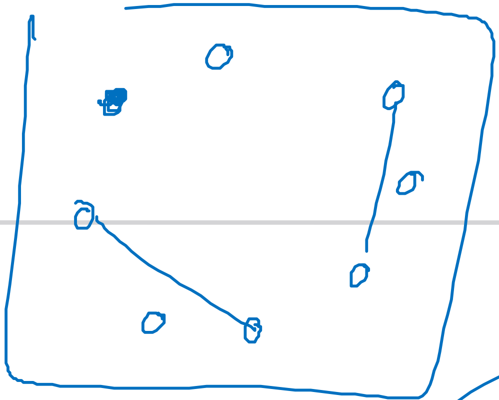
Epidemics in Static Random Graphs



- Random graph $G_{n,p}$
  - n nodes
  - Each directed edge occurs with independent probability p
- Expected indegre $\gamma = p\,(n-1)$
- How fast does an epidemic spread in $G_{n,p}$, if $\gamma \in O(1)$?
- Observation für n>2:
  - With probability $\geq 4^{-\gamma}$ and $\leq e^{-\gamma}$
    - a node has in-degree 0 and cannot be infected
    - a node has out-degree 0, and cannot infect others
- Implications:
  - Random (static) graph is not a suitable graph for epidemics

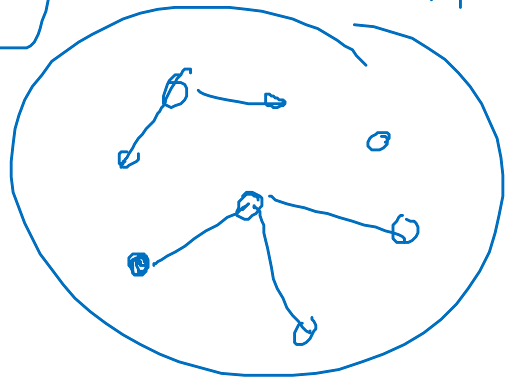
Erdős, Renyi

1960

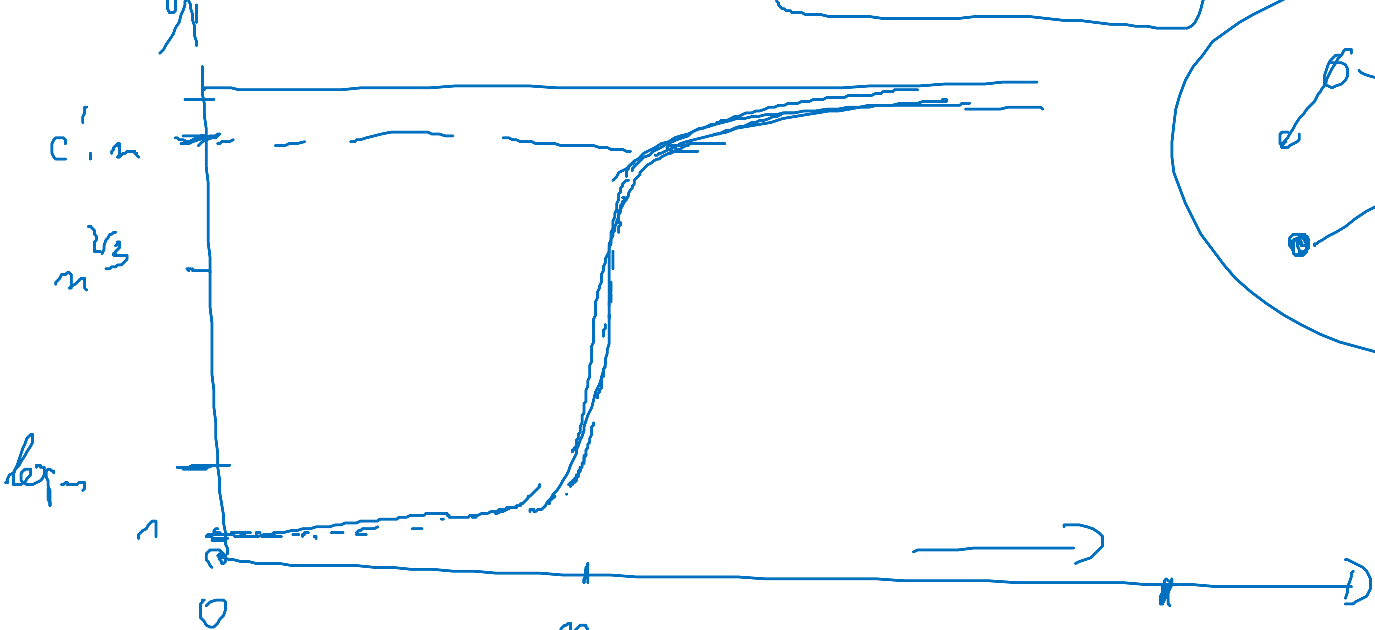


$n-1 \Big/ \boxed{\frac{n(n-1)}{2}}$

$\approx \frac{2}{n}$



\# connected neighbor



$n \cdot \log n$

$c \cdot \frac{n}{2}$

$n \cdot \log n$ # edges

$c > 1$