

FENÔMENO DO MUNDO PEQUENO

Redes Sociais E Econômicas

Prof. André Vignatti

“6 GRAUS DE SEPARAÇÃO” OU FENÔMENO DO MUNDO PEQUENO

as redes sociais têm **vários caminhos curtos**

esse fato se chama **fenômeno do mundo-pequeno** ou “6 graus de separação”

1ª experiência: feita por **Stanley Milgram:**

- **participantes** tentam **encaminhar carta** a um “alvo”
- o “alvo” tinha **nome, cidade, endereço**, etc... mas só podiam **enviar a carta para conhecidos**
- **~ 1/3** das cartas chegaram ao destino, em uma **média** de **6** passos

FENÔMENO DO MUNDO PEQUENO

dois fatos marcantes:

1. redes sociais têm **vários caminhos curtos**
2. pessoas sem “**mapa global**” são eficazes ao **encontrar tais caminhos curtos**

1º fato parece ser fácil de acreditar

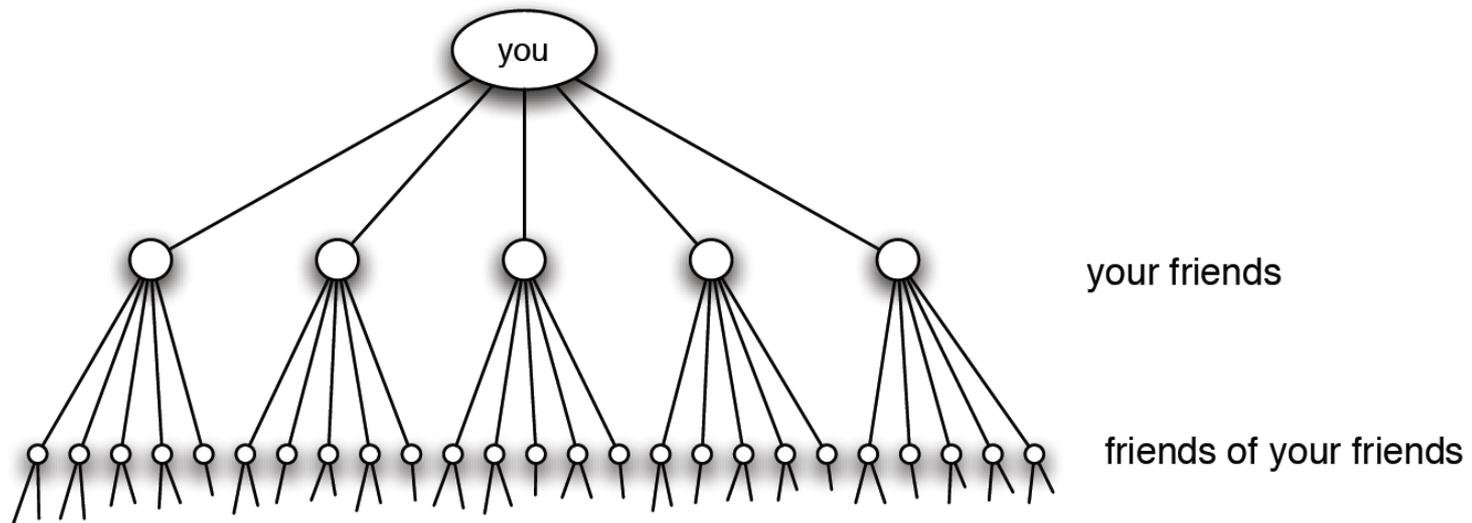
2º fato NÃO é tão fácil de acreditar:

- enviar **carta à pessoa** com CPF **123.456.789-0**, **daria errado**
- geralmente usa-se **informações geográficas e sociais**

ESTRUTURA E ALEATORIEDADE

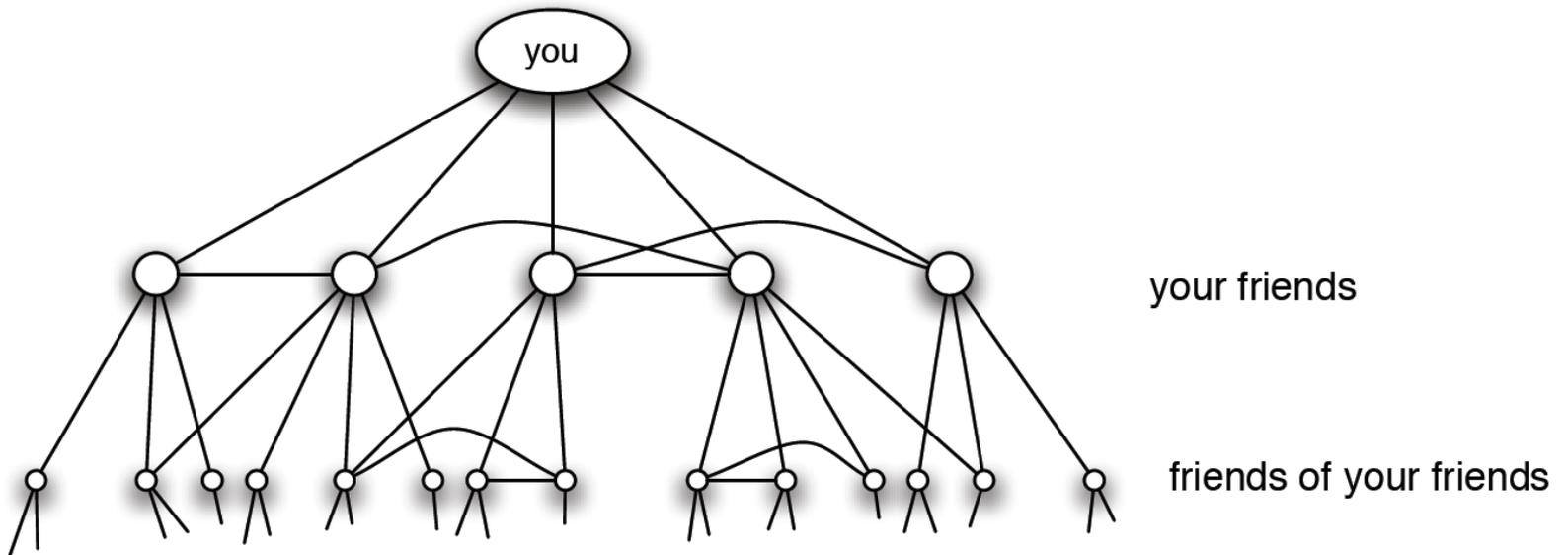
obter *modelos formais* para as duas características

suponha que você tem **100** amigos, cada um tem **100** amigos \Rightarrow em **duas etapas** que você se conecta a $100 \times 100 = 10.000$ pessoas



ESTRUTURA E ALEATORIEDADE

mas **fechamento triádico** limita o número de pessoas que você pode alcançar



assim, uma rede social tem **clusters**, aumentando o tamanho dos caminhos

O MODELO DE WATTS-STROGATZ

modelo: pessoas em **grade 2D** (um modo de modelar proximidade geográfica)

dois princípios básicos explicam caminhos curtos:

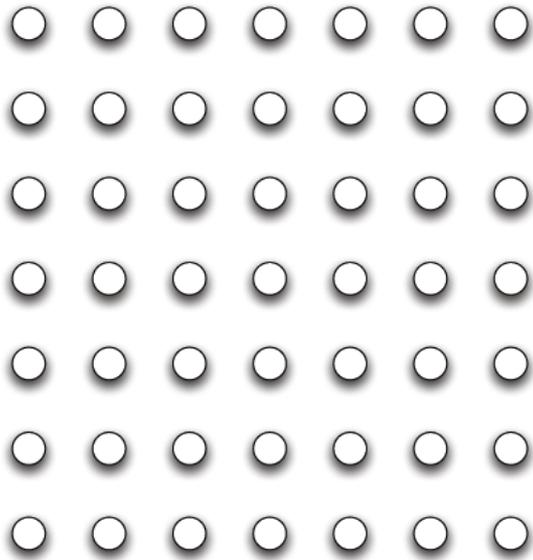
1. homofilia: nó tem link para outros num **raio** de r passos na grade

- cria **vários triângulos**

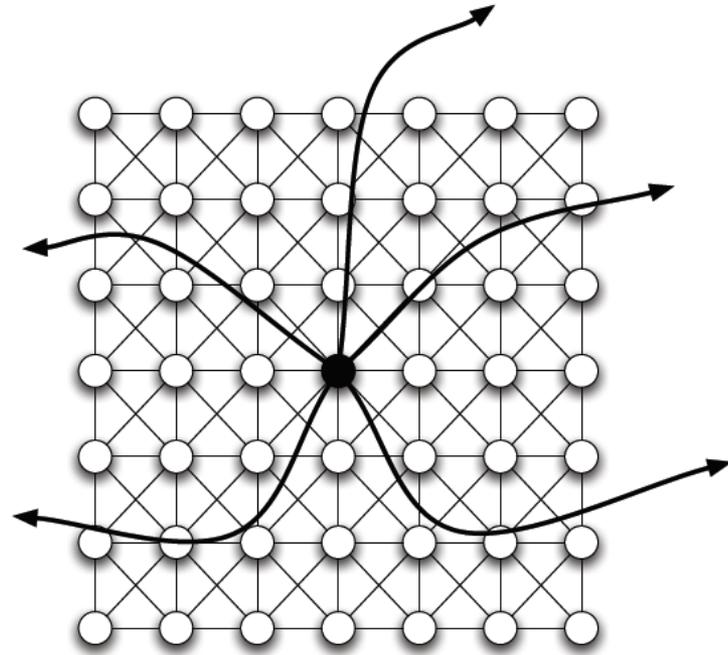
2. os laços fracos: nó tem link para k **outros nós aleatoriamente**

- atingem **nós distantes**

O MODELO DE WATTS-STROGATZ



(a) *Nodes arranged in a grid*



(b) *A network built from local structure and random edges*

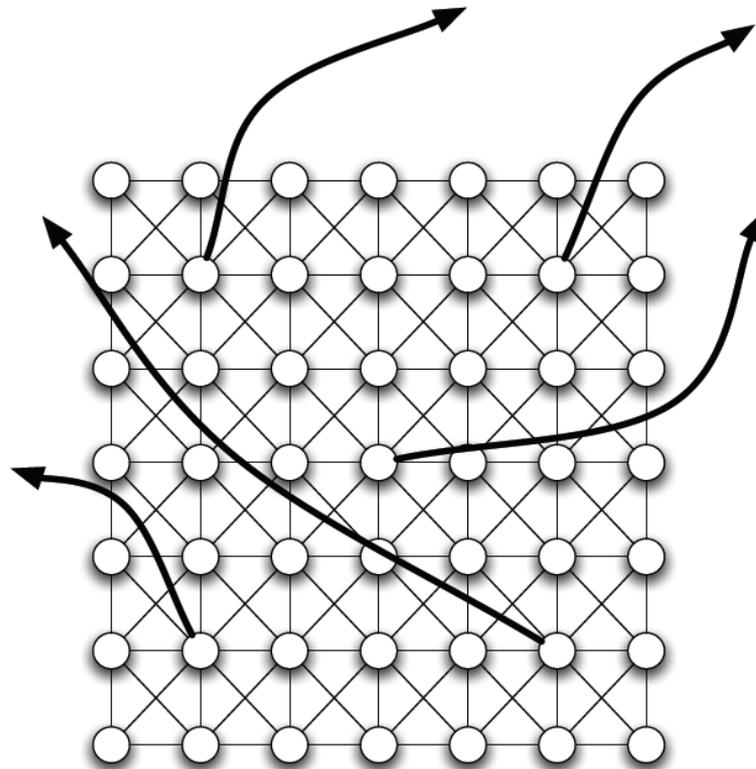
Teorema: o diâmetro esperado é $\theta(\log n)$

(não vamos ver a demonstração)

O MODELO DE WATTS-STROGATZ

na verdade **é necessário MUITO menos aleatoriedade** para manter os **caminhos curtos**:

- suponha que só é permitido **um único link aleatório** para cada k nós



O MODELO DE WATTS-STROGATZ

explicação: quadrado $k \times k$ (uma cidade) tem k links aleatórios

- considera-se o fenômeno a nível de cidades
- pode-se achar **caminhos curtos entre pares de cidades**
- dentro de uma cidade, usamos os links de proximidade (homofilia)

principal resultado do modelo de Watts-Strogatz:

pouca aleatoriedade torna o mundo pequeno

BUSCA DESCENTRALIZADA

2° fato do experimento de milgram: pessoas encontram caminhos curtos **sem “mapa” global**

modelo de Watts-Strogatz **não explica!**

- problema: **laços fracos** são *“muito aleatórios”*
- *links aleatórios* só levam **para muito longe**

Teorema: busca descentralizada no Watts-Strogatz usa **muitos passos**

novo modelo: **Kleinberg** - links levam a **distâncias aleatórias curtas, médias e longas**

MODELO PARA BUSCA DESCENTRALIZADA

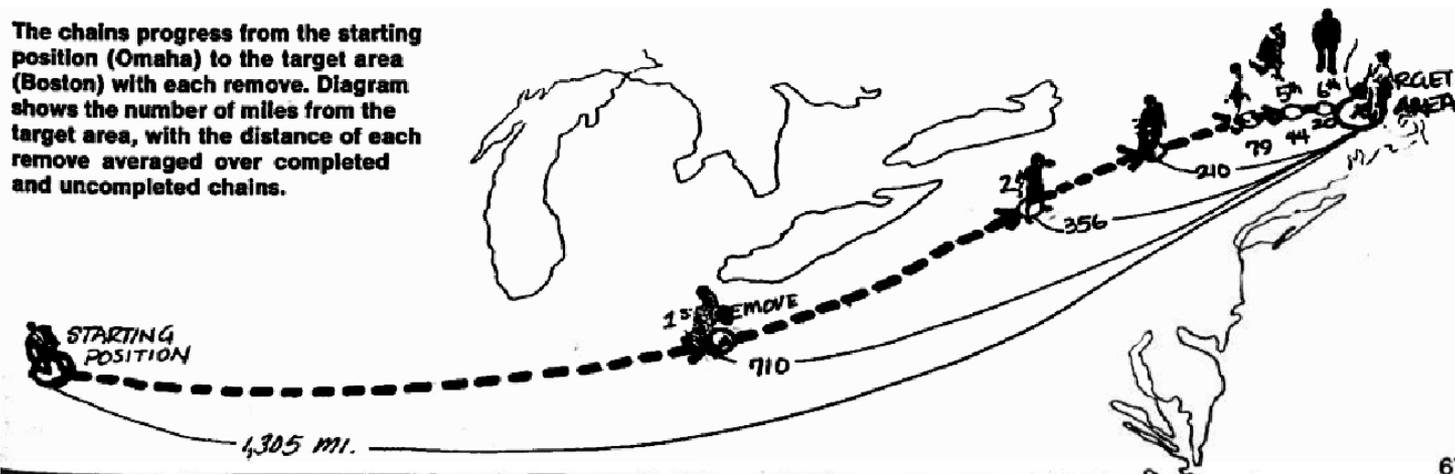
nó S deve enviar mensagem para nó t

S sabe apenas a **localização** de t na **grade**

mas S não conhece **nenhuma aresta de qualquer outro nó**

arestas aleatórias saindo de S , abrangendo **faixas intermediárias de escala**

The chains progress from the starting position (Omaha) to the target area (Boston) with each remove. Diagram shows the number of miles from the target area, with the distance of each remove averaged over completed and uncompleted chains.



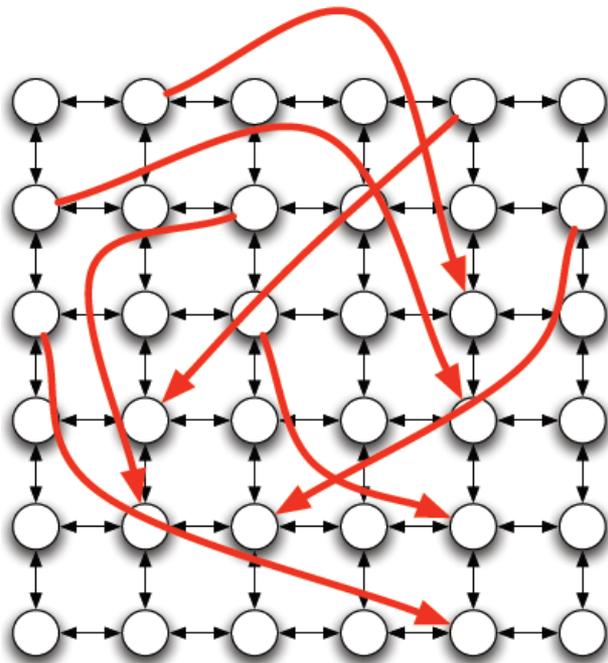
MODELAGEM DO PROCESSO DE BUSCA DESCENTRALIZADA

faixas intermediárias: expoente de clustering q :

- $d(v, w)$: número de passos entre v e w
- arestas aleatórias geradas com **probabilidade** $1/d(v, w)^q$
 - $q = 0$: links escolhidos **uniformemente de maneira aleatória**

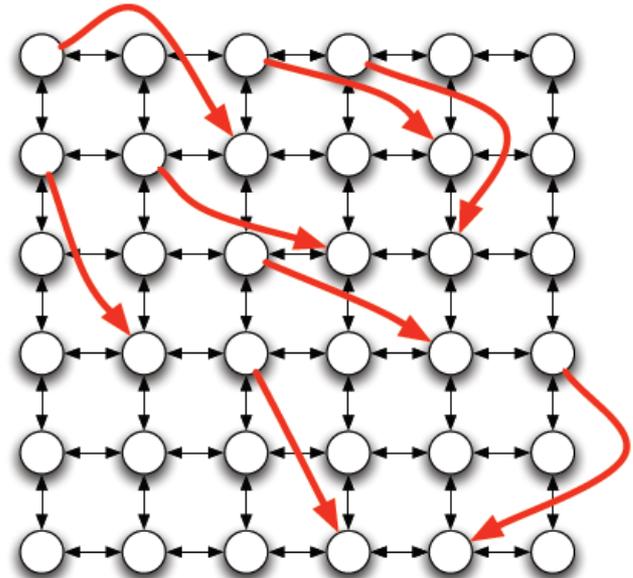
VARIANDO O EXPOENTE DE CLUSTERING

q **pequeno**: links de longo alcance são “**muito aleatórios**”: somente saltos longos



(a) *A small clustering exponent*

q **grande**: links de longo alcance “**não são aleatórios o suficiente**”: somente saltos curtos

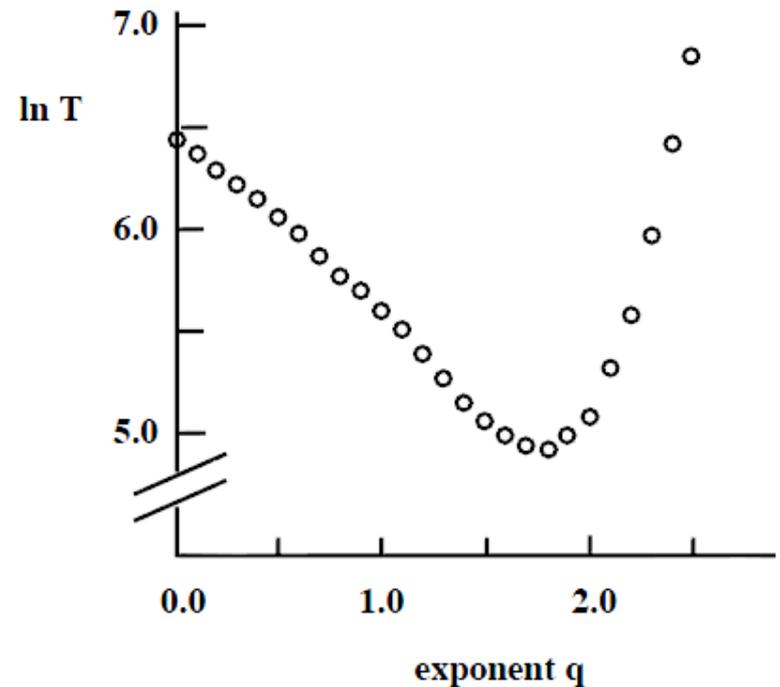


(b) *A large clustering exponent*

BUSCA DESCENTRALIZADA QUANDO $Q = 2$

qual q faz a **busca descentralizada rápida?**

experimentos: **mais eficiente quando $q = 2$** (links aleatórios seguem a distribuição do inverso do quadrado)



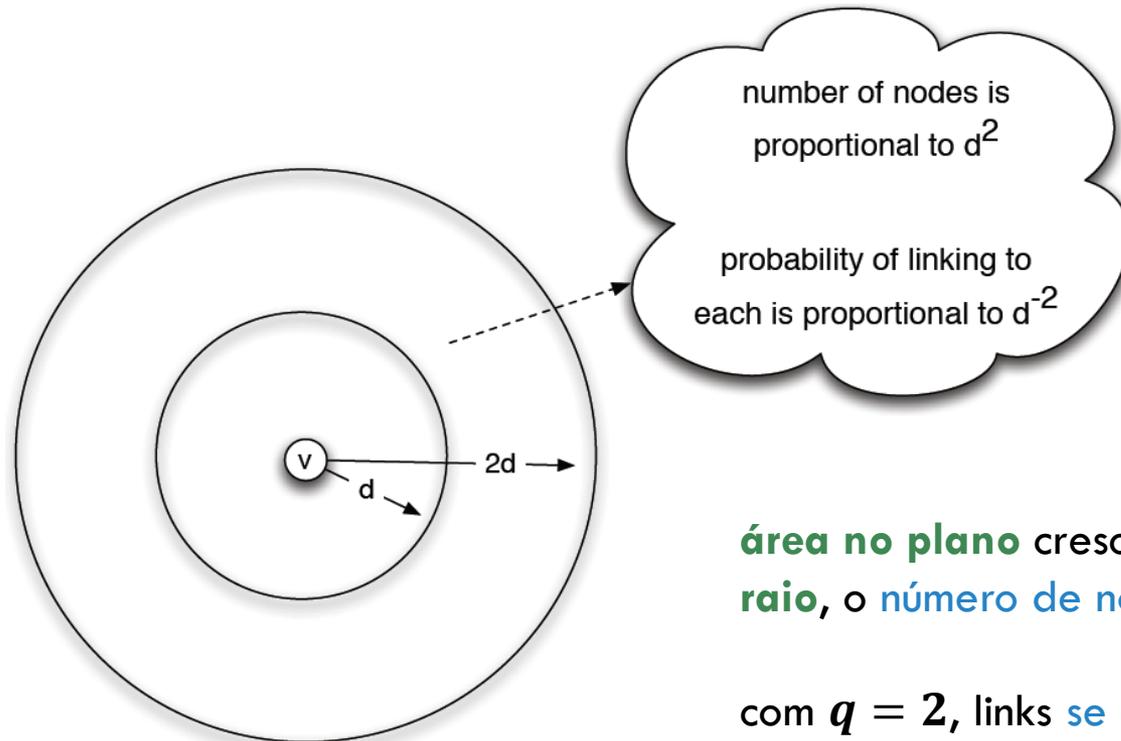
EXPLICAÇÃO DE $Q = 2$

o que há de **especial** no $q = 2$?

no mundo real, **mentalmente** organizamos distâncias em diferentes “**escalas de resolução**”:

- algo pode estar **no mundo, no país, no estado, na cidade, no quarteirão**
- ou seja, são **intervalos** de distância **cada vez maiores**: a distância de **2 – 4, 4 – 8, 8 – 16, ...**

EXPLICAÇÃO DE $Q = 2$



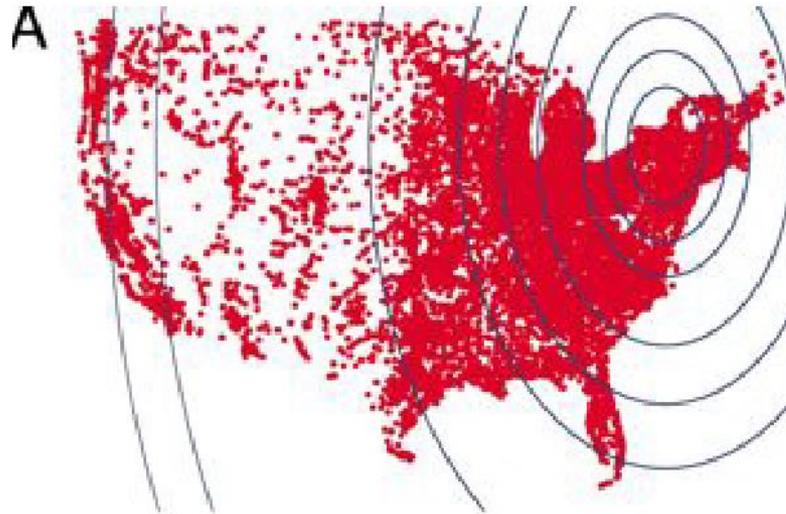
área no plano cresce com o **quadrado do raio**, o **número de nós** num grupo é $\sim d^2$

com $q = 2$, links **se espalham igualmente** entre **todos os grupos de escalas**

DADOS GEOGRÁFICOS SOBRE AMIZADE

análise empírica para verificar $q = 2$

análise de ~ 500 mil usuários e suas ligações no [LiveJournal](#) (só quem forneceu CEP)



problema: distribuição geográfica **não uniforme** (o modelo de $q = 2$ é útil somente para **pontos uniformemente distribuídos** no plano)

AMIZADE BASEADA EM RANKS

alternativa: criar links **usando seu rank** (posição relativa) ao invés de distância

- **elimina o problema** de **distribuições geográficas não uniformes**

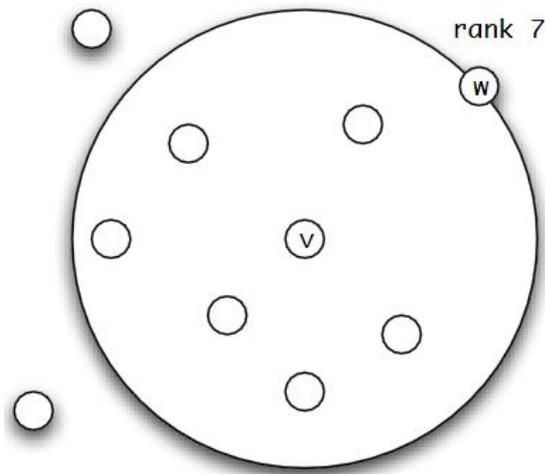
suponha que um nó v rankeia outros nós pela proximidade: o **ranking** de w é

$$\mathit{rank}_v(w) = i, \quad \text{se } w \text{ é o } i\text{-ésimo mais próximo } v$$

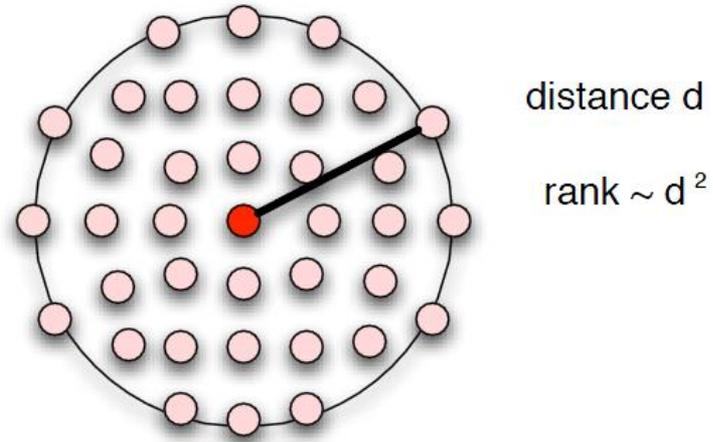
- ou seja, o nó **mais próximo** de v tem **rank 1**, o **segundo mais próximo** tem **rank 2**, etc...

AMIZADE BASEADA EM RANKS

o nó v cria um link aleatório escolhendo w com probabilidade $1/\text{rank}_v(w)^p$



(a) w is the 7th closest node to v .



(b) Rank-based friendship with uniform population density.

AMIZADE BASEADA EM RANKS

se $p = 1$ e população uniformemente distribuída $\rightarrow d^2$ nós mais próximos

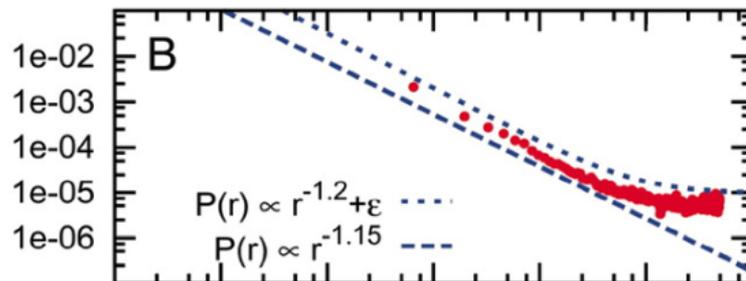
assim, $p = 1$ é **generalização** da distribuição do inverso do quadrado

propriedade: para construir uma rede que tem **busca descentralizada eficiente**, crie um link para cada nó com **probabilidade que é inversamente proporcional** ao **número de nós mais pertos**

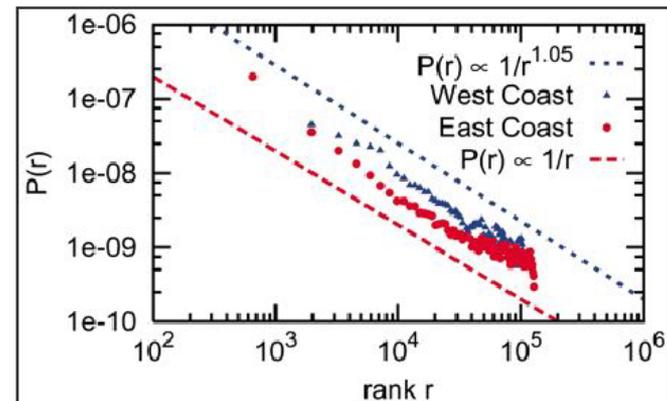
AMIZADE BASEADA EM RANKS

voltando ao **experimento do LiveJournal**:

- consideramos **pares de nós**, onde um atribui ao outro um rank igual a r
- qual a **fração f** desses pares que são **realmente amigos**?
- essa fração decresce proporcional a $1/r$? (ou seja, é uma power law?)



(a) Rank-based friendship on LiveJournal



(b) Rank-based friendship: East and West coasts

o mesmo para o **Facebook** (\sim rank 0,95)

FENÔMENOS DO MUNDO PEQUENO: SEQUÊNCIA DE CONCLUSÕES

1. experimento de Milgram: **(a)** parece que **há caminhos curtos** e **(b)** as pessoas sabem como **encontrá-los de forma eficaz**
2. construir **modelos matemáticos** para **(a)** e **(b)**
3. fazer uma **previsão com base nos modelos**: expoente de clustering **$q = 2$**
4. **validar esta previsão** usando dados reais de grandes redes sociais (*LiveJournal, Facebook*)