

Leonardo Aguiar do Amaral

**Algoritmos de Amostragem por Rejeição Dinâmica
Aplicados a Busca Descentralizada em Redes de
Mundo Pequeno com Topologia Fractal.**

Brasil

23 de Novembro de 2021

Leonardo Aguiar do Amaral

**Algoritmos de Amostragem por Rejeição Dinâmica Aplicados a
Busca Descentralizada em Redes de Mundo Pequeno com
Topologia Fractal.**

Tese apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências Exatas da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Doutor em Física Aplicada.

Universidade Federal do Espírito Santo – UFES
Programa de Pós-Graduação em Física

Orientador: Humberto Belich Jr

Brasil

23 de Novembro de 2021

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

A485a Amaral, Leonardo Aguiar do, 1981-
Algoritmos de amostragem por rejeição dinâmica aplicados a busca descentralizada em redes de mundo pequeno com topologia fractal / Leonardo Aguiar do Amaral. - 2021.
100 f. : il.

Orientador: Humberto Belich Junior.
Tese (Doutorado em Física) - Universidade Federal do Espírito Santo, Centro de Ciências Exatas.

1. Fractal. 2. Amostragem por Rejeição Dinâmica. 3. Small World. 4. Navegação em Redes. 5. Network Science. I. Belich Junior, Humberto. II. Universidade Federal do Espírito Santo. Centro de Ciências Exatas. III. Título.

CDU: 53



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

“Algoritmos de Amostragem por Rejeição Dinâmica Aplicados à Busca Descentralizada em Redes de Mundo Pequeno com Topologia Fractal”

Leonardo Aguiar do Amaral

Tese submetida ao Programa de Pós-Graduação em Física da Universidade Federal do Espírito Santo, por webconferência, meet.google.com/rbz-pkdx-cjt como requisito parcial para a obtenção do título de Doutor em Física.

Aprovada por:

Prof. Dr. Tales Costa Soares
(IF SUDESTE MG)

Prof. Dr. Wesley Spalenza
(IFES-Cariacica/ES)

Dr. Cássio Cecato Favarato
(IFES – São Mateus)

Prof. Dr. Humberto Belich Junior
(Orientador - PPGFis/UFES)

Prof. Dr. Antônio Canal Neto
(PPGFis/UFES)

Prof. Dr. Vinícius Cândido Mota
(PPGFis/UFES)

Vitória-ES, 23 de novembro de 2021.



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

PROTOCOLO DE ASSINATURA



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por
VINICIUS CANDIDO MOTA - SIAPE 1969881
Departamento de Física - DF/CCE
Em 17/12/2021 às 16:06

Para verificar as assinaturas e visualizar o documento original acesse o link:
<https://api.lepisma.ufes.br/arquivos-assinados/335267?tipoArquivo=O>

Este trabalho é dedicado a minha esposa Desiree e filhos, Arthur e Bianca.

Agradecimentos

À CAPES pelo apoio financeiro.

Ao Prof. Dr, Humberto Belich Jr pela orientação.

Ao José Carlos (secretário do PPGFis) pelo auxílio na solução de questões administrativas e por sempre estar disponível.

À minha esposa Desiree, pelo amor e companheirismo. Nos últimos anos passamos por muitas provações e tristezas com o adoecimento de nossa filha, hoje curada. Essa conquista tem um significado muito especial para nós.

Aos meus filhos, Arthur e Bianca, por darem sentido e propósito a minha vida.

À minha princesa (Bianca), por auxiliar na confecção da [Figura 31](#), escolhendo a cor rosa para a grade B_u .

À todos os meus familiares e amigos que não foram mencionados aqui, mas que sem eles esse trabalho não teria sido possível.

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (Capes) – Código de Financiamento 001.

*“O único lugar onde o sucesso vem
antes do trabalho é no dicionário.”
(Albert Einstein)*

Resumo

Nesta pesquisa estudamos a navegação de Kleinberg em redes de mundo pequeno através do algoritmo de amostragem por rejeição dinâmica, proposto por Mathieu e Comte. Ao contrário das técnicas convencionais, esta abordagem dispensa a utilização de condições de contorno periódicas que deformam a rede e a converte em toroide. Ao invés disso, utiliza-se máscaras de aceitação sobrepostas a rede quadrada estudada com o intuito de oferecer um critério de aceitação adicional para aquisição de links de longo alcance, sem contudo, violar os preceitos fundamentais que regem a distribuição estatística dessas conexões, resultando numa queda drástica da complexidade do algoritmo. O objetivo deste trabalho é estender a abordagem de Mathieu e Comte a redes com topologia fractal (Quadrado de Sierpinski), isso é feito através do desenvolvimento de uma rotina computacional auxiliar (*fractal_search*) que torna a geometria fractal detectável às máscaras de aceitação. Esta é uma etapa de grande relevância para a totalidade do processo no que diz respeito a complexidade do algoritmo, uma vez que o conhecimento acumulado em literatura na última década indica que a emergência de roteamento em redes fractais exige que estas possuam tamanhos extremamente elevados. Nossa rotina foi projetada para operar de forma recursiva e harmônica com a autossimilaridade fractal, o que lhe garante desempenho compatível com as tarefas que irá desempenhar durante a simulação. A abordagem adotada na condução desse trabalho se mostrou consistente com as previsões e resultados já consolidados e publicados em periódicos especializados na área de ciência de rede, tais como: a existência de um expoente de clusterização que minimiza o tempo de entrega, quando este assume valores idênticos ao da dimensão da rede analisada ($\alpha_{min} = d_f$), além do surgimento da proporcionalidade do tempo de entrega em relação ao logaritmo do comprimento da rede ($T \sim \ln^2 L$). Também ficou demonstrado que o número de realizações (R), utilizado para compor o espaço amostral das análises estatísticas da simulação, e o comprimento da rede (L), possuem interdependência e causam impacto expressivo no tempo de processamento. Finalmente, realizamos um teste de desempenho tomando como parâmetro comparativo o tempo de execução, em que ficou demonstrada a superioridade do método proposto em comparação aos métodos tradicionais.

Palavras-chave: Fractal. Amostragem por Rejeição Dinâmica. Small-World. Navegação em redes. Network Science.

Abstract

In this research, we study Kleinberg's navigation in small-world networks using the dynamic rejection sampling algorithm proposed by Mathieu and Comte. Unlike conventional techniques, this approach dispenses with the use of periodic boundary conditions that deform the lattice and convert it into a toroid. Instead, acceptance masks are used superimposed on the studied square network to offer an additional acceptance criterion for the acquisition of long-range links, without violating the fundamental precepts that govern the statistical distribution of these connections, resulting in a drastic drop in the complexity of the algorithm. This work aims to extend Mathieu and Comte's approach to networks with fractal topology (Sierpinski's Square), this is done through the development of an auxiliary computational routine (*fractal_search*) which makes fractal geometry detectable to acceptance masks. This is a step of great relevance for the entire process regarding the complexity of the algorithm since the knowledge accumulated in the literature in the last decade indicates that the emergence of routing in fractal networks requires that they have extremely large sizes. Our routine was designed to operate recursively and harmonically with fractal self-similarity, which guarantees a performance compatible with the tasks it will perform during the simulation. The approach adopted in conducting this work was consistent with the forecasts and results already consolidated and published in journals specialized in the field of network science, such as the existence of a clustering exponent that minimizes the delivery time, when it assumes values identical to the dimension of the analyzed network ($\alpha_{min} = d_f$), besides the appearance of the proportionality of the delivery time with the logarithm of the network size ($T \sim \ln^2 L$). It was also demonstrated that the number of realizations (R), used to compose the sampling space of the statistical analysis of the simulation, and the network size (L), have interdependence and have a significant impact on the processing time. Finally, we performed a performance test using the execution time as a comparative parameter, which demonstrated the superiority of the proposed method compared to traditional methods.

Keywords: Fractal. Sampling by Dynamic Rejection. Small-World. Browsing in networks. Network Science.

Lista de ilustrações

Figura 1 – O Problema das sete pontes	28
Figura 2 – A teoria de grafos aplicada ao problema das sete pontes	29
Figura 3 – Representação gráfica de grafos direcionados e não direcionados.	30
Figura 4 – Exemplo de grafos, subgrafos e suas respectivas notações matriciais.	31
Figura 5 – Coeficiente de clusterização: Avaliando a conectividade entre vizinhos.	34
Figura 6 – Estudo da aplicação da Assortatividade: As redes são classificadas como assortativas quando k_{nn} aumenta (Rede de colaboração científica), neutra quando não há correlação entre k_{nn} e k (Power Grid) e disassortativa quando k_{nn} diminui com k (Rede Metabólica).	36
Figura 7 – Exemplo de rede real formada através de processos aleatórios.	40
Figura 8 – A convergência da distribuição Binomial para Poisson ocorre para redes esparsas ($N \rightarrow \langle k \rangle$).	42
Figura 9 – Impossibilidade de distinção entre as distribuições Binomial e Poisson para redes com o mesmo valor de $\langle k \rangle$ e diferentes valores de N	43
Figura 10 – Exemplo ilustrativo de redes reais que não são modeladas por processos totalmente aleatórios.	44
Figura 11 – Similaridade: Exemplos de transformações	48
Figura 12 – Exemplo de Auto-Similaridade na natureza	48
Figura 13 – Auto-Afinidade para fractais anisotrópicos	49
Figura 14 – Triângulo de Sierpinski: Formação do fractal com crescimento interior ao inicializador	50
Figura 15 – Fractais determinísticos e estatísticos para o quadrado de Sierpinski.	51
Figura 16 – Relações de dimensionalidade	52
Figura 17 – Determinação da Dimensão fractal	53
Figura 18 – A densidade em fractais.	54
Figura 19 – Método de Watts e Strogatz para geração de redes de mundo pequeno com $N = 15$ e $m = 2$: (a) Rede totalmente regular com $p = 0$ (b) Rede com algumas ligações refeitas com probabilidade $p = 0,2$	56
Figura 20 – Abordagem de Watts e Strogatz: Inclusão de atalhos através do aumento da aleatoriedade.	57

Figura 21 – Simulações produzidas com o modelo de Watts e Strogatz para redes de mundo pequeno.	57
Figura 22 – Experimento de Milgram: A distância euclidiana entre o portador da mensagem e o alvo diminui a metade conforme a mensagem se aproxima do alvo.	58
Figura 23 – Exemplo de vizinhança e contatos de longa distância segundo proposições de Kleinberg.	60
Figura 24 – Estratificação da trajetória no modelo de Kleinberg.	61
Figura 25 – Condições de contorno periódicas aplicadas em métodos tradicionais de simulação	62
Figura 26 – Navegação de Kleinberg em redes de pequeno mundo com $d = 2$, $p = 1$ e $q = 1$. Prova computacional de que a rede se torna navegável em $\alpha = 2$	62
Figura 27 – Criação de link de longo alcance entre sítios do quadrado de sierpinski	63
Figura 28 – Exemplo de máscara de aceitação para um rede 2D.	64
Figura 29 – Análise da vizinhança de um vértice para criação das máscaras de aceitação.	65
Figura 30 – Aplicação de <i>Fractal Search</i> : Redução da geração do fractal $n = 3, 2, 1$ nas pastes (a), (b) e (c) respectivamente. Como o ponto v pintado de vermelho não se encontra na região central de nenhuma das redes G_3, G_2, G_1 a rotina informa que v é navegável.	70
Figura 31 – Amostragem por rejeição dinâmica em redes com geometria fractal.	72
Figura 32 – Navegação de Kleinberg em redes de mundo pequeno com geometria fractal utilizando a amostragem por rejeição dinâmica acoplada a <i>Fractal Search</i> : (a) Tempo de entrega para redes $L = 3^n$ tal que $n \in \Theta$: Em todos os casos $\alpha_{min} \neq d_f$ sem violar o roteamento de Kleinberg. Comportamento sugere que α_{min} converge para d_f com o aumento de n . (b) Consistência da simulação com a previsão $T \sim (LnN)^2$ apresentada na Equação 4.4.	74
Figura 33 – Ajuste parabólico de curva para obtenção de α_{min}	76
Figura 34 – Obtenção da dimensão fractal d_f pelo método <i>Finite-Size Corrections</i>	78
Figura 35 – A convergência $d \rightarrow d_f$ é lenta.	78
Figura 36 – Análise do tempo de execução	79
Figura 37 – Distribuição de Poisson do grau k	95

Lista de tabelas

Tabela 1 – Redes reais e sua inadequação ao modelo de redes totalmente aleatórias.	45
Tabela 2 – Decaimento da densidade fractal para o Triângulo de Sierpinski.	54
Tabela 3 – Métricas utilizadas por Watts e Strogatz para identificar e simular redes de mundo pequeno artificialmente.	56

Lista de símbolos

G	Grafo
N	Número de vértices de um grafo
M	Número de arestas de um grafo
V	Conjunto de N vértices de um grafo G
A	Conjunto de M arestas
X	Matriz de adjacência
D	Diâmetro do grafo
ρ	Densidade de um grafo
E	Matriz de distância
$\langle L \rangle$	Comprimento médio do caminho mais curto em grafos
k_i	Grau de um Vértice
$P(k)$	Distribuição de Graus de um Vértice
$\langle k \rangle$	Grau Médio de um vértice da rede
C_i	Coefficiente de Clusterização do vértice i
$\langle C \rangle$	Coefficiente de agregação médio
$k_v(k)$	Grau médio dos vizinhos de um nó com grau k
λ	Coefficiente de assortatividade
C_Δ	Coefficiente de Agrupamento Global
α	Expoente de clusterização
T	Tempo de entrega

Sumário

Introdução	23
I Referenciais Teóricos	25
1 Network Science	27
1.1 Introdução	27
1.2 A criação da Teoria de Grafos	28
1.2.1 Conceitos Básicos e Definições	29
1.2.1.1 Grafos e Representação Gráfica	29
1.2.1.2 Grafos e Representação Matricial	30
1.2.1.3 Densidade de Grafos	31
1.2.1.4 Medidas de Distância em Grafos	32
1.2.1.5 Grau de um Vértice	32
1.2.1.6 Entropia	33
1.2.1.7 Coeficiente de Clusterização	33
1.2.1.8 Assortatividade de redes	35
1.2.1.9 Medidas de Centralidade	36
2 Redes Aleatórias	39
2.1 Introdução	39
2.2 O modelo de redes aleatórias	39
2.2.1 Redes completamente aleatórias	39
2.2.2 Redes Reais	43
3 Fractais	47
3.1 Auto-Similaridade	47
3.2 Auto-Afinidade	49
3.3 A construção de Fractais	49
3.4 A Dimensão Fractal	51
3.5 A densidade de Fractais	53
4 Algoritmos de geração de redes artificiais	55
4.1 Redes de mundo pequeno (Small-World)	55
4.2 O experimento de Milgram	58
4.3 Redes de Kleinberg	59
4.3.1 Construção do modelo	59
4.3.2 Tempo de Entrega	60
4.4 Navegação em redes de mundo pequeno fractal	63
4.5 Amostragem por rejeição dinâmica	64

II Proposições e Simulações	67
5 Máscaras de aceitação fractal	69
5.1 Introdução	69
5.2 Algoritmo de busca fractal (Fractal Search)	69
5.3 Máscaras de aceitação fractal	71
6 Resultados numéricos e discussões	73
6.1 Parâmetros adotados na simulação e regime computacional	73
6.2 Resultados numéricos	73
6.2.1 A existência de um α que minimiza T e consistência com o modelo de Kleinberg	73
6.2.2 A convergência $d \rightarrow d_f$	74
6.2.3 Análise de desempenho	77
7 Conclusão	81
Referências	83
Apêndices	91
APÊNDICE A Relação entre as Distribuições Binomial e Poisson	93
APÊNDICE B Valores máximo e mínimo do grau médio $\langle k \rangle$ de um vértice	95
APÊNDICE C Modelo de Kleinberg: Tempo de entrega para $\alpha = 2$	97

Introdução

A simples prática de estabelecer novos vínculos de amizade, revela um aspecto curioso e emergente das redes sociais. Em situações como estas, frequentemente encontramos membros antigos em nossa rede de contatos que já apresentavam algum tipo de conexão com a pessoa de nossa amizade recém adquirida, ainda que de forma abstrata. Isso ocorre porque redes sociais são repletas de caminhos curtos que permitem navegar por toda a sua extensão com um número relativamente pequeno de intermediários, chamamos isso de *Fenômeno de Mundo Pequeno* ou de *Seis Graus de Separação*. O primeiro estudo empírico de relevância sobre esse fenômeno, foi conduzido em (TRAVERS; MILGRAM, 1967). A pesquisa consistiu no envio de cartas para uma pessoa-alvo previamente escolhida, os remetentes por sua vez, eram escolhidos de forma aleatória e possuíam apenas algumas informações elementares sobre o alvo como nome, endereço e profissão. Evidentemente, o envio direto até o alvo não era permitido, ao invés disso, o remetente só poderia enviar a carta para um de seus contatos que estivesse mais próximo do alvo e assim sucessivamente até que o alvo fosse atingido. Surpreendentemente, entre todos os caminhos possíveis, a carta trafegava pela cadeia de conexões mais curtas na maioria das vezes, apesar de o portador dispor apenas de informações locais. Em ciências da computação, algoritmos que simulam tais navegações são chamados de *Busca Descentralizada* ou de *Algoritmo Guloso*. Outra terminologia que surge durante esse estudo é o chamado *Tempo de Entrega T*, que se refere ao número de intermediários que portaram a carta ao longo de sua trajetória até o destino. A partir do trabalho de Milgram, inúmeros esforços foram despendidos visando a compreensão do fenômeno de mundo pequeno. Tradicionalmente, a teoria de grafos é a abordagem preferida por pesquisadores da área e aparentemente é a mais eficaz. Nesse ramo da matemática, os modelos de redes totalmente aleatórias proposto por Erdos e Renyi (ERDÖS; RÉNYI, 1960a; ERDÖS; RÉNYI, 1959b) e por Gilbert (GILBERT, 1959) foram de grande importância para a construção de um modelo satisfatório de rede, pois também apresentam cadeias curtas de conexão ao longo da rede. Mesmo não sendo consistentes com o experimento de Milgram, esses modelos foram de grande serventia para o entendimento e concepção de modelos mais aprimorados de redes, como o modelo de Watts-Strogatz (WATTS; STROGATZ, 1998) que consiste em simular a rede entre os extremos, totalmente regular e totalmente aleatório, através de um parâmetro de controle que rege o grau de aleatoriedade com que as conexões são refeitas. O trabalho de Watts-Strogatz mostrou

que a aleatoriedade é o ingrediente que favorece a emergência de caminhos curtos (atalhos) mas não explicava porque esses caminhos ocorriam na maioria das realizações, o simples fato desses caminhos existirem não é garantia de que serão encontrados. Essa questão foi abordada por Kleinberg (KLEINBERG, 1999; KLEINBERG, 2000), ao contrário dos modelos anteriores, ele propôs que a criação de atalho não deveria ser uniformemente distribuída, mas sim criada com uma distribuição de probabilidade consistente com as observações de Milgram, ou seja, a cada passo que a carta está mais próxima do alvo sua distância é reduzida para cerca de metade da distância anterior. Como consequência natural desta imposição, atalhos mais curtos são mais prováveis, em detrimento dos mais longos. Alguns anos depois, o modelo de Kleinberg foi estendido para navegação em redes com geometria fractal (ROBERSON; BEN-AVRAHAM, 2006) e todos os modelos desenvolvidos até então foram baseados num método de modelar a rede denominado condições de contorno periódicas, que tem como principal efeito a conversão da rede quadrada em um toroide com o intuito de remover interferências causadas por efeitos de borda. Alternativamente, Mathieu e Comte (MATHIEU; COMTE, 2020) trabalharam numa implementação do algoritmo de busca descentralizada de Kleinberg em redes 2D que não requer a aplicação das condições de contorno periódicas. Este método ficou conhecido como *Amostragem por Rejeição Dinâmica* e consiste no uso de máscaras de aceitação para desenhar links de longo alcance, o que melhorou consideravelmente o desempenho do algoritmo.

Nesta pesquisa, estudamos a navegação de Kleinberg em redes de pequeno mundo e propomos a extensão do método Mathieu e Comte para simular o tempo de entrega esperado T em redes com geometria fractal conhecidas como Tapete de Sierpinski. Validamos nossos resultados utilizando técnicas analíticas e computacionais semelhantes às já catalogadas na literatura. Ao longo do texto, utilizamos linguagem e ilustrações acessíveis a todos os públicos, com o intuito de tornar o conteúdo menos árido a leitores não especializados no tema. No capítulo 1 discutimos as ferramentas básicas de abordagem no âmbito da ciência de rede e teoria de grafos. Nos capítulos 2 e 3 examinamos o processo de evolução dos modelos de rede mais relevantes ao nosso interesse de pesquisa, bem como os conceitos básicos sobre a geometria fractal. No capítulo 4 apresentamos e discutimos os principais algoritmos de busca descentralizada já publicados em periódicos. Nossa contribuição científica consta nos capítulos 5 e 6, onde descrevemos a metodologia de nossa abordagem e apresentamos os resultados obtidos via simulação.

Parte I

Referenciais Teóricos

Network Science

1.1 Introdução

Neste capítulo serão discutidos os principais tópicos correlatos a representação e modelagem das conexões entre pares de elementos de um mesmo sistema. Ao longo das seções será evidenciado que essa representação é feita através da *Ciência de Rede*¹, e que o aporte matemático necessário para a modelagem é fornecido pela *teoria de grafos*.

"I think the next century will be the century of complexity."

Stephen Hawking

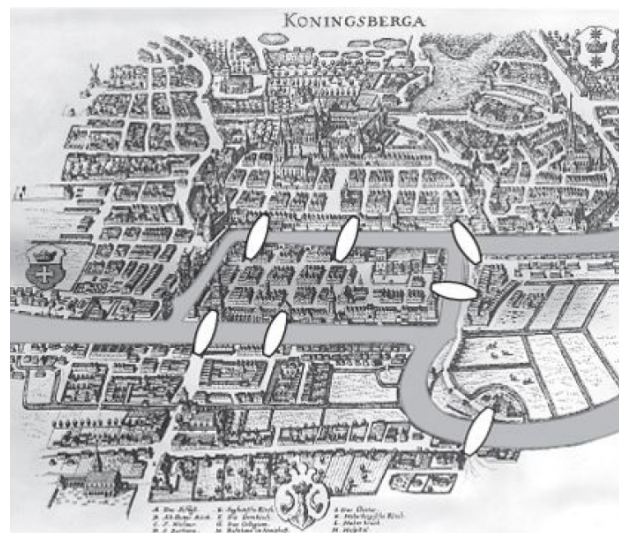
Sistemas com elevado número de componentes que potencialmente podem interagir entre si são denominados *Sistemas Complexos*. Convenientemente esta terminologia possui relação com complexidade descritiva das interações entre os elementos desse sistema, o que torna inviável qualquer tentativa de confeccionar um modelo matemático-analítico através de métodos convencionais. Vamos mostrar que a ciência de rede é uma ferramenta poderosa no estudo de problemas com esse grau de dificuldade e que os principais resultados de interesse perpassam por uma análise topológica das conexões aliada a uma abordagem reducionista do problema até que um modelo satisfatório seja alcançado.

¹ Network Science

1.2 A criação da Teoria de Grafos: O problema das sete pontes de Königsberg

Em 1735, na cidade prussiana de Königsberg, haviam duas ilhas separadas por um rio chamado Pregel (atual rio Pregolia), e a conexão entre elas e o continente se dava através de sete pontes, conforme ilustrado na [Figura 1](#). Moradores se perguntavam se seria possível percorrer toda a cidade passando pelas sete pontes, uma única vez em cada. Evidentemente ninguém conseguiu encontrar tal caminho, o que estimulava as pessoas a continuarem tentando já que não existia prova de que ele não existia. Desde então, o desafio de encontrar uma solução definitiva para esta questão ficou conhecido como o *problema das sete pontes de Königsberg* ([WEISSTEIN, 2002](#); [BULA](#); [NIKIEL](#); [TYMCHATYN, 1994](#)).

Figura 1 – O Problema das sete pontes (cidade de Königsberg em 1735).

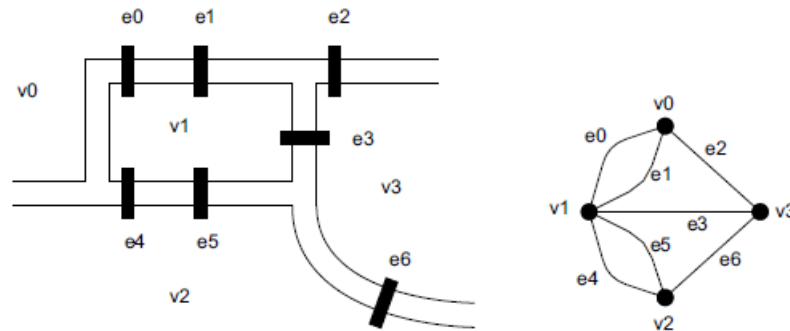


Fonte: ([COHEN; HAVLIN, 2010](#))

Nesta mesma época Leonhard Euler passou a estudar o problema e sua primeira contribuição foi mostrar que os comprimentos das pontes não eram determinantes mais sim como as porções de terra se conectavam, deixando claro que a ferramenta matemática adequada ao caso é o que conhecemos hoje como teoria de grafos. A [Figura 2](#) mostra o grafo que representa o problema proposto ([HOPKINS; WILSON, 2004](#); [GRIBKOVSKAIA; SR; LAPORTE, 2007](#)).

Euler provou que o caminho procurado no problema das sete pontes era análogo ao de se percorrer um caminho no grafo. Foi mostrado que para um caminho genérico, o

Figura 2 – Na teoria de grafos aplicada ao problema das sete pontes, cada porção de terra e ponte são associados a um vértice e a uma aresta, respectivamente.



Fonte: Newman, Barabási e Watts (2006)

*grau dos nós*² de saída e chegada poderiam ser ímpares (apenas chegada ou apenas saída) e o grau dos nós intermediários deveriam ser par (uma chegada e uma saída), como o caso estudado tratava de um caminho fechado ficou evidenciado que todos os nós do grafo deveriam possuir grau par. Por inspeção, podemos perceber que no grafo do problema existem nós com número ímpar de ligações, isso fez com que Euler concluísse que o caminho procurado não existia (EULER, 1741; GONDRAN; MINOUX; VAJDA, 1984).

1.2.1 Conceitos Básicos e Definições

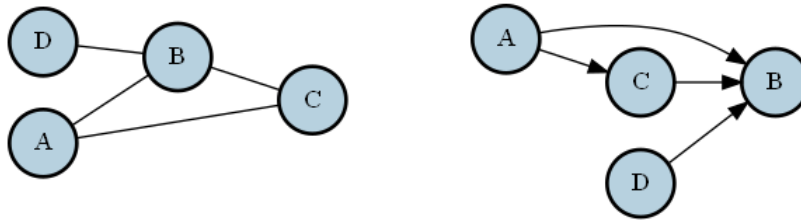
Nesta subseção discutiremos as principais métricas utilizadas na caracterização de grafos. Os tópicos abordados aqui farão parte do aporte teórico utilizado na construção do modelo de rede que estudaremos neste trabalho.

1.2.1.1 Grafos e Representação Gráfica

Segundo Barabási (2013) para entender um sistema complexo de forma efetiva é preciso conhecer o mecanismo de interação existente entre seus componentes. No linguajar da teoria de grafos, os constituintes do sistema são chamados de *nós* ou *vértices* e as eventuais conexões entre eles de *links* ou *arestas* (MARCUS, 2008; BOLLOBÁS, 2013). Dessa forma, um sistema de N elementos pode ser representado por um grafo $G = (V, A)$, onde V é um conjunto de N vértices que identifica os constituintes do sistema e A outro conjunto que identifica o número M de arestas bem como sua configuração topológica. Dizemos que a ordem do grafo e seu tamanho são respectivamente N e M (TRUDEAU, 2013; GROSS; YELLEN; ZHANG, 2013).

² Termo utilizado na teoria de redes que diz respeito ao Número de Ligações de um nó. Maiores detalhes na seção 1.2.1

Figura 3 – Representação gráfica de grafos direcionados e não direcionados.



Fonte: Autor

Os elementos de A são representados por um par de informações que dizem respeito aos vértices conectados, dessa forma, dois nós i e j podem ser conectados através de uma aresta $(i, j) \in A$ sendo chamados de nós *vizinhos* ou *adjacentes*. Quando $(i, j) = (j, i)$ dizemos que o grafo é não direcionado e quando $(i, j) \neq (j, i)$ ele é chamado de direcionado. Neste trabalho vamos lidar com grafos sem autoconexão, isto é, um nó do grafo não pode conectar-se a si mesmo. A [Figura 3](#) mostra uma representação gráfica de dois grafos como exemplo, a esquerda temos um grafo não direcionado com $V = \{A, B, C, D\}$ e $A = \{(A, B), (A, C), (B, C), (B, D)\}$ e a direita um grafo direcionado com $V = \{A, B, C, D\}$ e $A = \{(A, B), (A, C), (C, B), (D, B)\}$. Podemos determinar o número máximo de conexões em grafos não direcionados através da combinação do número de nós N , tomados dois a dois, o que nos leva a

$$C_{N,2} = \binom{N}{2} = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}. \quad (1.1)$$

Evidentemente que para grafos direcionados temos o dobro de conexões, ou seja, $N(N-1)$ ([TRINAJSTIC, 2018](#); [FOULDS, 2012](#); [WALLIS, 2010](#)). O escopo do nosso trabalho possui pouca ou nenhuma aderência aos grafos direcionados, por isso vamos omiti-lo nos desenvolvimentos teóricos que se seguem.

1.2.1.2 Grafos e Representação Matricial

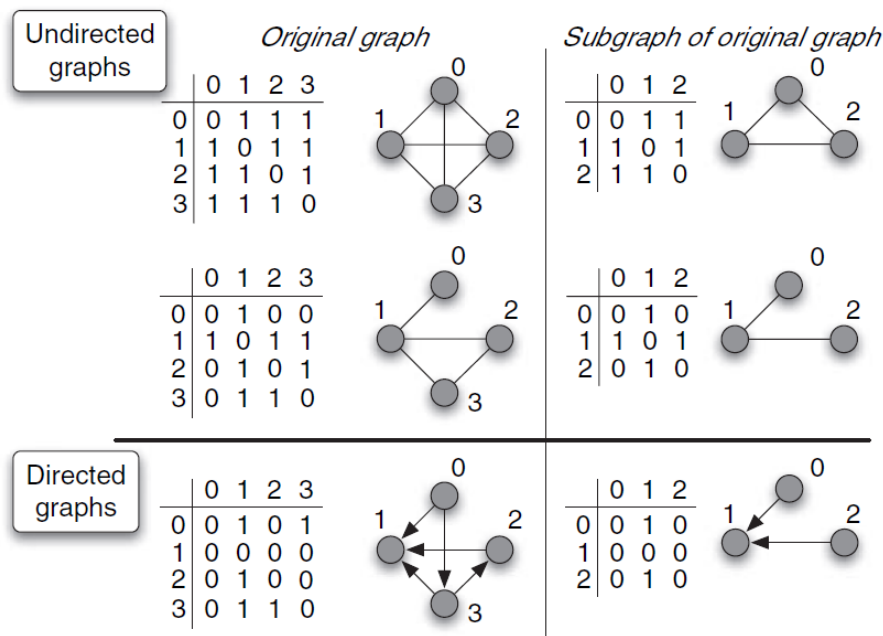
Uma forma de representação de grafos mais rigorosa em termos analíticos pode ser feita através de matrizes, e a matriz relacionada é denominada *Matriz de Adjacência* X . Ela possui tamanho $N \times N$ e seus elementos x_{ij} são binários de modo que $x_{ij} = 1$ se $(i, j) \in A$ e $x_{ij} = 0 \notin A$. Como optamos por excluir qualquer autoconexão temos que $x_{ii} = x_{jj} = 0$; note que para grafos não direcionados X é sempre simétrica.

Apenas como caráter informativo, apresentamos a seguir a matriz de adjacência do grafo não direcionado ilustrado na [Figura 3](#).

$$X = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Na [Figura 4](#) oferecemos mais exemplos de grafos e seus respectivos subgrafos, em ambas os casos direcionado e não direcionado.

Figura 4 – Exemplo de grafos, subgrafos e suas respectivas notações matriciais.



Fonte: [Barrat, Barthelemy e Vespignani \(2008\)](#)

1.2.1.3 Densidade de Grafos

Outra métrica de grande importância na caracterização de grafos é a *Densidade* ρ , ela mede a razão entre o número de conexões existentes e o número máximo de conexões possíveis, um grafo é chamado de denso quando $\rho \simeq 1$ ou esparsos quando $0 < \rho \ll 1$. Dessa forma, para os grafos não direcionados tratados aqui temos que

$$\rho = \frac{2M}{N(N-1)}. \quad (1.2)$$

De acordo com [Monteiro \(2010\)](#), a representação matricial não é adequada quando se trata de grafos esparsos devido a grande quantidade de zeros que devem ser armazenados.

Ele sugere agrupar os dados utilizando listas, em que cada nó é associado a uma *Lista de Adjacência*.

1.2.1.4 Medidas de Distância em Grafos

Neste trabalho, vamos considerar exclusivamente os casos onde não haja vértices desconectados, para isso, é necessário um número mínimo $N - 1$ de conexões (GIBBONS, 1985; BARNES; HARARY, 1983). Um grafo nessas condições é dito conexo e será chamado de desconexo em caso contrário. Conseqüentemente, para grafos não direcionados e conexos, sempre haverá um caminho (cadeia de conexões) que conecta dois vértices quaisquer i e j do grafo. Dessa forma, o número de arestas l_{ij} que constituem a cadeia mais curta de conexões entre dois nós quaisquer é calculada para todas as combinações de pares possíveis e armazenada numa matriz denominada *Matriz de Distância E* (RAHMAN et al., 2017; JOHN; ALLAN, 1995). Como ilustração, apresentamos abaixo a matrizes de distância referente ao grafo não direcionado mostrado na Figura 3.

$$E = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{pmatrix}$$

Sistemas complexos são modelados tipicamente por grafos de ordem N bastante elevada, dentro de um contexto estocástico (STEEN, 2010; HSU; LIN, 2008). Por isso, uma estimativa do comprimento médio do caminho mais curto $\langle L \rangle$ é mais um dado que colabora para a análise da rede estudada.

$$\langle L \rangle = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N l_{ij} \quad (1.3)$$

Dentre todos os caminhos mais curtos l_{ij} da distribuição, o mais longo é denominado *Diâmetro do Grafo D*.

1.2.1.5 Grau de um Vértice

O *Grau de um Vértice* k_i é o número de arestas que ele possui e é obtido através da matriz de adjacência pela seguinte expressão:

$$k_i = \sum_{j=1}^N x_{ij}. \quad (1.4)$$

Uma característica observada por Euler é que a soma dos graus de todos os nós deve ser

$$\sum_{i=1}^N k_i = 2M. \quad (1.5)$$

Uma breve inspeção dos dados permite concluir que cada uma das arestas é contada duas vezes, justificando assim o fator dois na [Equação 1.5](#). O elevado número da ordem N dos grafos abordados nesse trabalho, nos encoraja a considerar alguns aspectos estatísticos da situação ([VASUDEV, 2006](#); [READ, 2014](#)). Ao longo da rede, o grau dos vértices seguem uma distribuição conhecida simplesmente como *Distribuição de Graus* $P(k)$ que pode ser interpretada como a fração de nós com grau k ou a probabilidade de um nó, escolhido aleatoriamente, possuir um grau k ([PHILLIPS; SCHWANGHART; HECKMANN, 2015](#); [SPIELMAN, 2007](#)). Conseqüentemente, o número médio de arestas por vértice, denominado *Gráu Médio* $\langle k \rangle$ de um grafo, é escrito como

$$\langle k \rangle = \sum kP(k). \quad (1.6)$$

1.2.1.6 Entropia

A entropia é uma medida da heterogeneidade da conectividade entre os nós grafo

$$H = - \sum P(k) \log_2(P(k)). \quad (1.7)$$

1.2.1.7 Coeficiente de Clusterização

O *Coeficiente de Clusterização* C_i é talvez uma das métricas aplicadas a teoria de grafos de maior relevância para o estudo que vamos propor aqui. Introduzida originalmente por [Watts e Strogatz \(1998\)](#), C_i é a razão entre o número a_i de conexões efetivas existentes entre os primeiros vizinhos e o total de ligações possíveis entre eles. Podemos interpretar C_i como o grau de interação entre os vizinho de um nó.

$$C_i = \frac{a_i}{\binom{N}{2}} = \frac{2a_i}{k_i(k_i - 1)}. \quad (1.8)$$

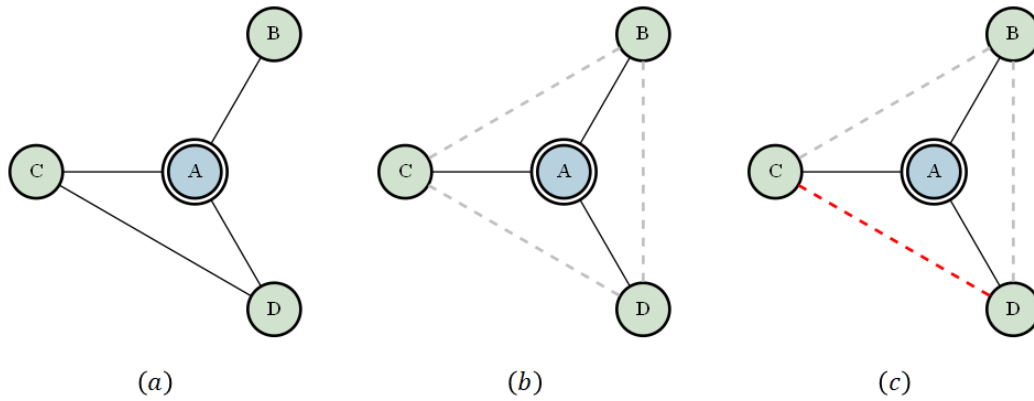
É possível verificar que a_i pode ser obtido através da matriz de adjacência correlacionada X ([SCHANK; WAGNER, 2005](#); [GERHARDT; LEMKE; CORSO, 2006](#)) conforme abaixo:

$$a_i = \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^N x_{ij} x_{jl} x_{li}. \quad (1.9)$$

Na Figura 5 apresentamos um exemplo aplicado, a parte (a) mostra um pequeno grafo onde o vértice A (Azul) terá seus vizinhos (Verde) estudados em termos de conectividade. Na parte (b) destacamos com linhas tracejadas na cor cinza todas as combinações $C_{3,2}$ de ligações possíveis entre os três vizinhos de A . Em (c) vemos que dentre todas as possibilidades existe apenas uma única conexão entre os vizinhos, especificamente entre os vértices C e D . Destacamos essa aresta com uma linha vermelha tracejada. Dessa forma temos

$$C_A = \frac{a_A}{\binom{N}{2}} = \frac{2 \cdot 1}{3(3-1)} = \frac{1}{3}.$$

Figura 5 – Coeficiente de clusterização: Avaliando a conectividade entre vizinhos.



Fonte: Autor

Para uma rede de N nós, o *coeficiente de clusterização médio* $\langle C \rangle$ é

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (1.10)$$

Em uma abordagem estatística $\langle C \rangle$ pode ser considerada como sendo a probabilidade de encontrar dois vértices, que são vizinhos de um terceiro, ligados entre si. Alternativamente, $\langle C \rangle$ pode ser substituído por outra métrica similar, comumente encontrada na literatura como *Coeficiente de Agrupamento Global*, cuja representação é feita por C_Δ . Ela consiste em medir o número total de triângulos fechados em uma rede, ou seja, o número de subgrafos de três vértices totalmente conectados. Uma tripla conectada é outro subgrafo de três vértices, porém conectados por apenas duas arestas de modo que seu diâmetro seja igual a dois,

$$C_\Delta = \frac{3 \times (\text{Número de Triângulos})}{\text{Número de Triplas Conectadas}}. \quad (1.11)$$

Por inspeção, podemos constatar que o termo a_i na [Equação 1.8](#) é o número total de triângulos nos quais o nó i participa. Note que $\langle C \rangle$ e C_Δ não são quantidades equivalentes ([BARABÁSI, 2013](#)), porém elas fornecem valores aproximados para um grande número de redes ([BOLLOBÁS; RIORDAN, 2003](#)).

1.2.1.8 Assortatividade

A assortatividade é uma métrica utilizada na medição da tendência de vértices portadores de características similares estarem conectados. Isso pode ser bastante útil na análise de redes sociais, onde se deseja analisar a similaridade entre pessoas da mesma classe social, nacionalidade, religião, faixa etária, etc.

Na última década, pesquisadores ([PASTOR-SATORRAS; VÁZQUEZ; VESPIGNANI, 2001](#)) propuseram uma forma de caracterizar a *assortatividade de grau* com o intuito de medir a correlação entre os graus dos nós em uma rede. Ficou estabelecido que uma rede é assortativa quando nós com grau considerado pequeno tendem a se conectar a nós com grau comparativamente pequeno. Por outro lado, quando essa tendência ocorre com nós portadores de elevado grau, dizemos que a rede é disassortativa.

Dessa forma temos $k_{v,i}$ como a soma de todos os graus dos vizinhos do i -ésimo nó comparado ao grau desse próprio nó definido como

$$k_{v,i} = \frac{1}{k_i} \sum_{j=1}^N x_{ij} k_j. \quad (1.12)$$

De um ponto de vista mais abrangente, podemos definir o grau médio dos vizinhos de um nó que apresenta grau k da seguinte forma

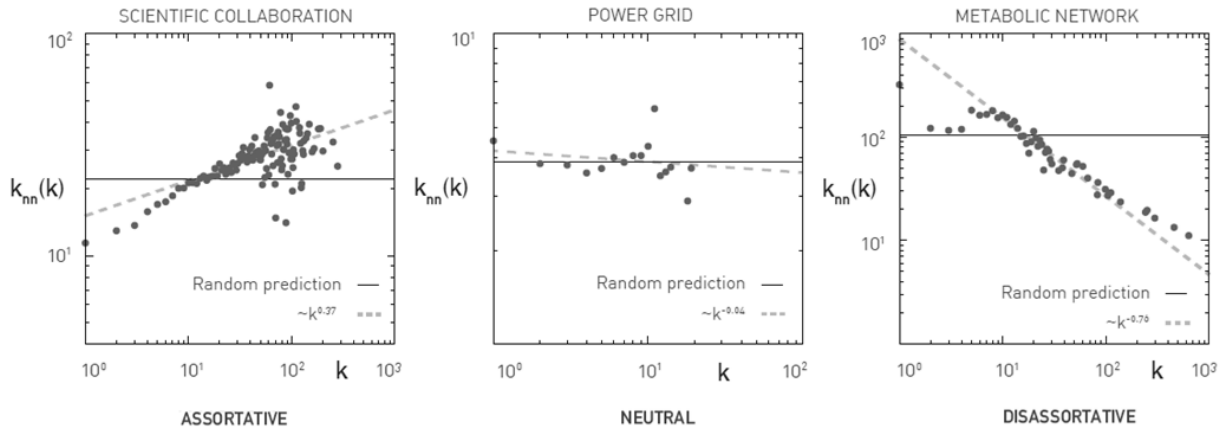
$$k_v(k) = \frac{1}{N_k} \sum_{i/k_i=k} k_{v,i}. \quad (1.13)$$

Na [Equação 1.13](#), N_k é o número de vértices portadores de grau igual a k e o somatório é realizado considerando apenas esses nós. Nota-se que em uma rede assortativa, $k_v(k)$ é diretamente proporcional a k , evidentemente se essa relação for inversa então a rede é disassortativa.

Outras propostas metodológicas de avaliação da assortatividade surgiram com o tempo, o trabalho de [Newman \(2002\)](#) aborda a questão propondo um coeficiente de assortatividade λ ,

$$\lambda \equiv \frac{\frac{1}{M} \sum_{i=1}^M j_i l_i - \left(\frac{1}{2M} \sum_{i=1}^M (j_i + l_i) \right)^2}{\frac{1}{2M} \sum_{i=1}^M (j_i^2 + l_i^2) - \left(\frac{1}{2M} \sum_{i=1}^M (j_i + l_i) \right)^2}. \quad (1.14)$$

Figura 6 – Estudo da aplicação da Assortatividade: As redes são classificadas como assortativas quando k_{nn} aumenta (Rede de colaboração científica), neutra quando não há correlação entre k_{nn} e k (Power Grid) e disassortativa quando k_{nn} diminui com k (Rede Metabólica).



Fonte: (BARABÁSI, 2013) - Adaptada

Na Equação 1.14, j_i e l_i denotam os graus dos vértices conectados através da i -ésimo link. O coeficiente λ tem como característica apresentar $\lambda > 0$ para redes assortativas e $\lambda < 0$ para redes disassortativas. Em particular, redes perfeitamente assortativas apresentam $\lambda = 1$ e redes perfeitamente disassortativas $\lambda = -1$. Evidentemente, para o caso onde não há correlação alguma entre a existência de arestas e os graus dos vértices que elas conectam, temos $\lambda = 0$. A Figura 6 mostra um breve estudo da assortatividade aplicado a algumas das redes de amplo conhecimento da área.

1.2.1.9 Medidas de Centralidade

Frequentemente, a análise integral de uma rede requer a identificação dos nós de maior relevância; aqueles que, caso sejam removidos, causariam grandes transformações na rede original (BORGATTI; EVERETT, 2006; ESTRADA; BODIN, 2008). Chamamos isso de centralidade de um vértice e ao longo do tempo os cientistas acumularam diversas maneiras de estimá-la. Neste trabalho apresentaremos quatro delas (MASON; VERWOERD, 2007).

A *centralidade de grau* $C_g(i)$ de um nó i é definida como a razão entre o grau desse nó e o maior grau que potencialmente ele possa ter

$$C_g(i) \equiv \frac{k_i}{N-1}. \quad (1.15)$$

A *centralidade de proximidade* $C_p(i)$ de um nó i é a razão entre os $N-1$ nós

restantes na rede e o somatório dos caminhos mais curtos entre i os demais nós da rede, tendo a expressão,

$$C_p(i) \equiv \frac{N-1}{\sum_{j=1}^N l_{ij}}. \quad (1.16)$$

Proposta originalmente por [Freeman \(1977\)](#), a *centralidade de intermediação* $C_{in}(i)$ do nó i é a razão entre o número de caminhos mais curtos $\sigma_{jl}(i)$ entre dois outros nós j e l que passam por i e o número total de caminhos existentes σ_{jl} entre j e l . A expressão é normalizada através de um fator que leva em consideração o número máximo de vezes em que o nó i pode surgir nesses caminhos, para um grafo não direcionado esse número é $\binom{N}{2}$, assim:

$$C_{in}(i) \equiv \frac{2}{(N-1)(N-2)} \sum_{j=1}^{N-1} \sum_{l=j+1}^N \frac{\sigma_{jl}(i)}{\sigma_{jl}}. \quad (1.17)$$

Apontamos a abordagem de [Bonacich \(1987\)](#) como a quarta e última metodologia de medição da centralidade. O procedimento de cálculo deve levar em conta a centralidade de todos os outros vértices aos quais o vértice estudado está conectado. Dessa forma:

$$v_i = \frac{1}{\Gamma} \sum_{j=1}^N x_{ij} v_j. \quad (1.18)$$

Sendo Γ um fator de proporcionalidade que manipulado adequadamente pode converter a [Equação 1.18](#) numa equação de autovalor que envolve a matriz de adjacência conforme abaixo

$$X \vec{v} = \Gamma \vec{v}. \quad (1.19)$$

Na [Equação 1.19](#), temos que Γ é autovalor e \vec{v} autovetor da matriz de adjacência X .

Dessa forma, a *centralidade de autovalor* $C_a(i)$ de um vértice i como a i -ésima componente do autovetor \vec{v} que se relaciona com o maior autovalor de X ([IRWIN; HUGHES, 1992](#)).

Redes Aleatórias

2.1 Introdução

Neste capítulo iremos apresentar os principais modelos teóricos aplicados a ciência de redes desenvolvidos nas últimas décadas e discutiremos sua evolução até o modelo aplicado ao objeto de estudo deste trabalho.

2.2 O modelo de redes aleatórias

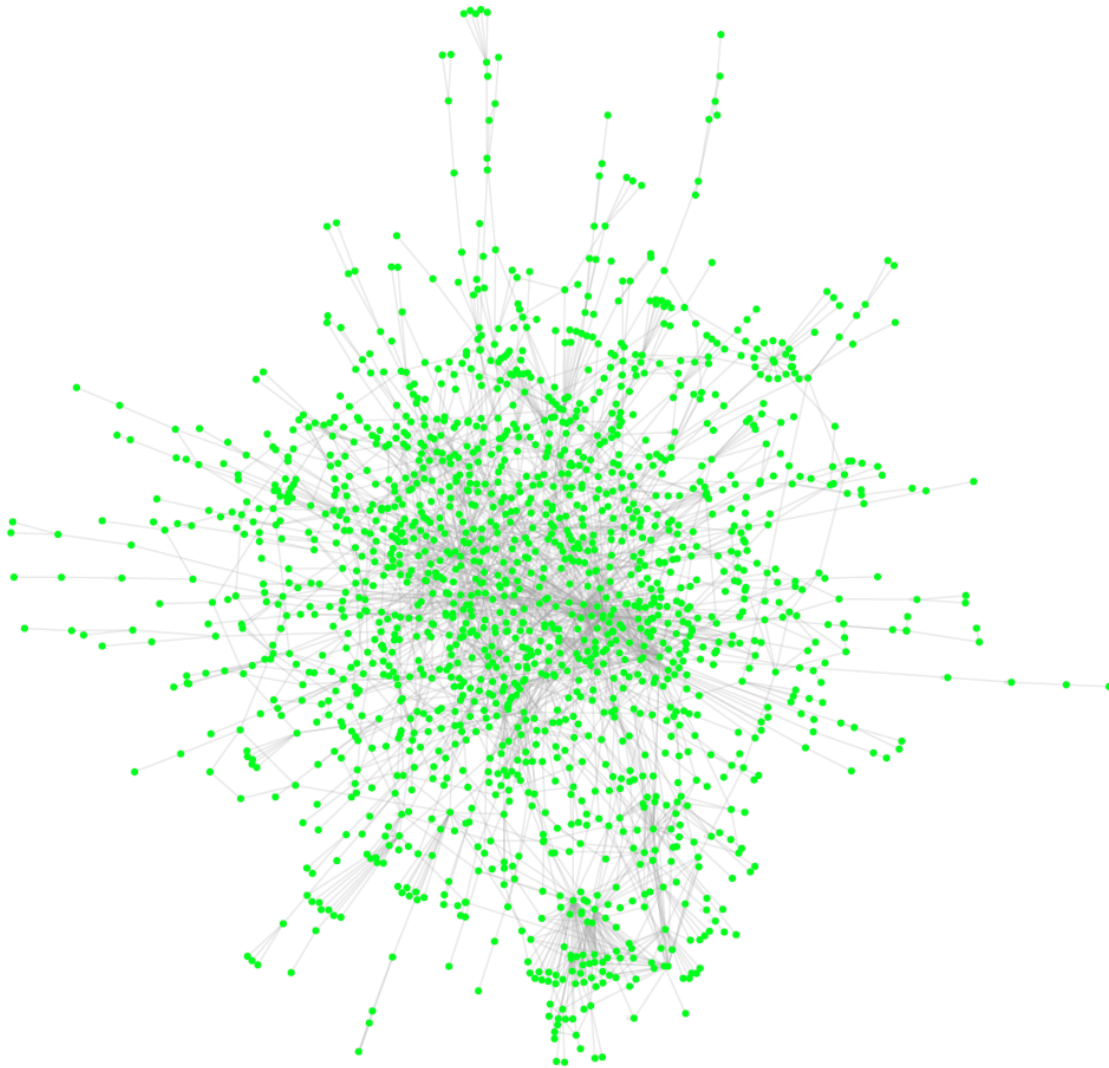
Um dos objetivos da ciência de rede é produzir modelos que representem as redes reais em todas as suas características e complexidade. Dessa forma, o desafio de qualquer modelo é determinar como as ligações serão estabelecidas.

A [Figura 7](#) ilustra uma rede de interação de proteínas de fermento em que cada vértice corresponde a uma proteína, evidentemente suas ligações foram obtidas através de técnicas experimentais. Por inspeção, acredita-se que a rede em questão possa ter sido construída através de processos aleatórios. De fato, nas seções que se seguem, mostraremos que a aleatoriedade ocupa um papel de grande importância no desenvolvimento e evolução da ciência de rede.

2.2.1 Redes completamente aleatórias

Modelos de redes totalmente aleatórias são aqueles em que as conexões são feitas de forma aleatória numa distribuição uniforme, ou seja, todos os nós possuem a mesma probabilidade de estarem conectados. Existem dois modelos similares atualmente aceitos, um proposto por E.N. Gilbert ([GILBERT, 1959](#)) e outro por P. Erdos e A. Rényi em

Figura 7 – Rede de interação de proteínas presente no fermento. Acredita-se que sua formação tenha se dado através de processos aleatórios.



Fonte: (BARABÁSI, 2013) - Adaptada

uma série de artigos (ERDÖS; RÉNYI, 1959a; ERDÖS; RÉNYI, 1960b; ERDÖS; RÉNYI, 1961a; ERDÖS; RÉNYI, 1961b; ERDÖS; RÉNYI, 1963; ERDÖS; RÉNYI, 1966a; ERDÖS; RÉNYI, 1966b; ERDÖS; RÉNYI, 1968).

No primeiro modelo, considera-se uma rede com N vértices em que a combinação demonstrada na equação [Equação 1.1](#) fornece o número total de arestas possíveis $N(N - 1)/2$. Assume-se que cada uma delas é criada com uma probabilidade p e não é criada com probabilidade $1 - p$. Dessa forma, parte-se de um grafo totalmente conectado de onde se removem arestas. Nesse modelo o número médio de conexões é $pN(N - 1)/2$ e os parâmetros N e p são fixos.

Já no segundo modelo, N vértices são conectados por M arestas tomadas ao acaso

entre as $N(N-1)/2$ possíveis. Os parâmetros N e M são fixos, e quando $M = pN(N-1)/2$ atinge-se a equivalência com o modelo de Gilbert.

Para redes geradas através dessas metodologias esperamos uma ligeira diferença a cada realização, mesmo mantendo os parâmetros N e p fixos. Isso ocorre devido a flutuações associadas ao número de links, por essa razão é de grande interesse determinar quantos links se espera obter em cada realização particular. Cada sorteio realizado para decidir sobre a existência de uma aresta é um ensaio de Bernoulli. Consequentemente, para testarmos todos os nós, vários ensaios serão necessários. Isso evidencia que a distribuição do número de links deve obedecer uma distribuição binomial.

Dessa forma, a probabilidade $p(L)$ de L tentativas produzirem conexões entre os $N(N-1)/2$ pares de nós é p^L e a probabilidade das tentativas restantes que não resultam em ligações é $(1-p)^{N(N-1)/2-L}$. A contagem do número de maneiras distintas que as ligações podem ser realizadas é dada pela combinação $\binom{\frac{N(N-1)}{2}}{L}$, sendo assim a distribuição binomial correlata é

$$p(L) = \binom{\frac{N(N-1)}{2}}{L} p^L (1-p)^{N(N-1)/2-L}. \quad (2.1)$$

Com valor esperado de

$$\langle L \rangle = \sum_{L=0}^{N(N-1)/2} L p_L = p \frac{N(N-1)}{2}. \quad (2.2)$$

A [Equação 2.2](#) nos permite concluir que o número máximo de vértices disponíveis para conexão é

$$L_{\max} = N(N-1)/2. \quad (2.3)$$

e o grau médio $\langle k \rangle$

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1). \quad (2.4)$$

Note que para grandes redes onde $N \gg 1$ a [Equação 2.4](#) pode ser aproximada para

$$\langle k \rangle \simeq pN. \quad (2.5)$$

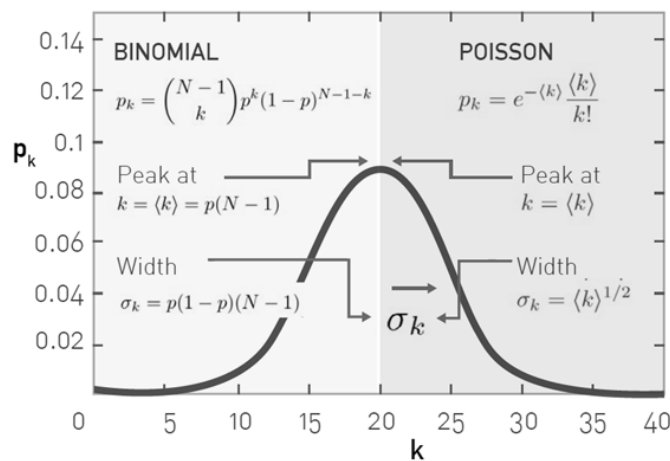
Por construção, a distribuição de graus $P(k)$ correspondente à probabilidade de um vértice escolhido ao acaso possuir k conexões também é binomial, conforme expresso na [Equação 2.6](#).

$$p(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (2.6)$$

Em termos estatísticos, destacamos o fato de que a distribuição binomial se aproxima da distribuição de Poisson na maioria das redes aleatória utilizadas no contexto da ciência de rede, ou seja, para elevados valores de N . No apêndice A apresentamos uma prova rigorosa dessa convergência, que resultou na Equação 2.7. Já na Figura 8, ilustramos um quadro comparativo das principais características descritivas de ambas as distribuições. Dessa forma, a versão de Poisson para a distribuição fica sendo

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (2.7)$$

Figura 8 – A convergência da distribuição Binomial para Poisson ocorre para redes esparsas ($N \rightarrow \langle k \rangle$).



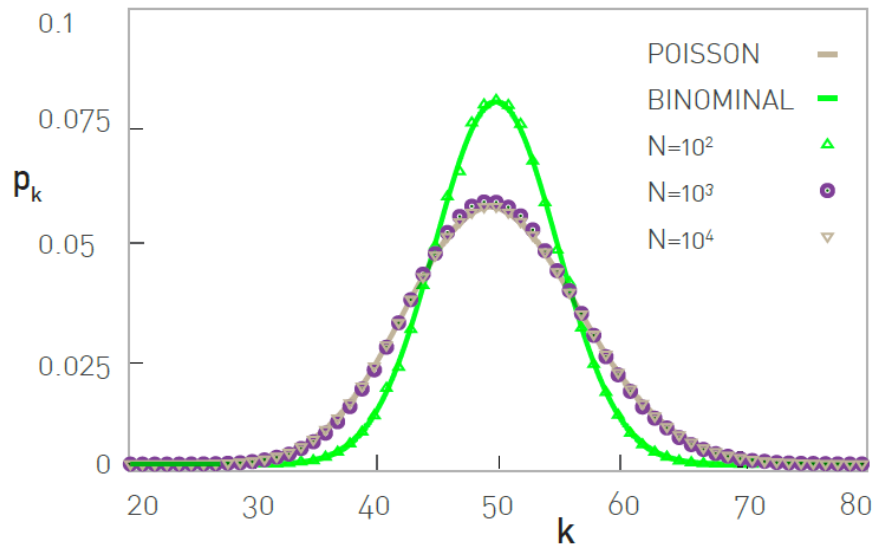
Fonte: (BARABÁSI, 2013) - Adaptada

Note que em ambos os casos o valor esperado $\langle k \rangle$ é o mais provável. De acordo com a Equação 2.5, um aumento de p ocasiona um aumento de $\langle k \rangle$, que por sua vez desloca a posição do pico do gráfico para a direita tornando a rede mais densamente conectada. A dispersão é controlada por p na distribuição binomial e por $\langle k \rangle$ na distribuição de Poisson, em ambos os casos, quanto mais densa for a rede maior será a sua dispersão.

A utilização da distribuição de Poisson apresenta algumas vantagens no que diz respeito a simplificação dos parâmetros oriundos da estatística descritiva da rede, tais como $\langle k \rangle$, $\langle k^2 \rangle$ e σ . No entanto, o fato de a distribuição não depender diretamente de N torna impossível distinguir a distribuição entre duas redes com o mesmo $\langle k \rangle$ e com N diferentes. Isso pode ser observado na Figura 9, onde é possível perceber a influência do tamanho (N) da rede sobre a forma da distribuição. Quando a rede é pequena (10^2)

os pontos se ajustam na distribuição binomial, enquanto que para valores mais elevados (10^3 e 10^4) o melhor ajuste ocorre na distribuição de Poisson. Note que em todos os casos $\langle k \rangle = 50$.

Figura 9 – Distribuição de k para redes aleatórias com valores de $N = 10^2, 10^3$ e 10^4 . As linhas verde e cinza são respectivamente as distribuições Binomial e Poisson. Em todos os casos $\langle k \rangle = 50$. Uma rede pequena (10^2) necessariamente segue a distribuição Binomial já as redes maiores (10^3 e 10^4) convergem para Poisson.



Fonte: (BARABÁSI, 2013)

Nessas condições, interpretamos $\langle k \rangle$ como o número médio de vértices vizinhos onde a distância entre cada um deles é uma unidade de modo que $\langle k \rangle \langle L \rangle \simeq N$. Isso nos permite fazer uma estimativa do comprimento médio do caminho mais curto

$$\langle L \rangle = \frac{\log N}{\log \langle k \rangle}. \quad (2.8)$$

2.2.2 Redes Reais

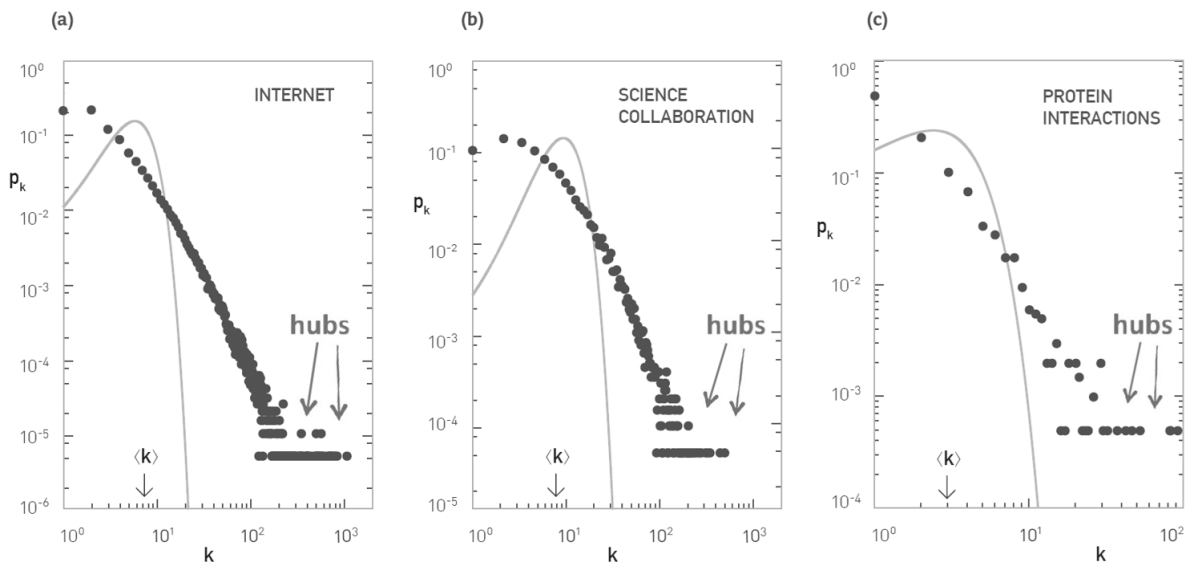
Nesta seção examinaremos algumas redes reais e a efetividade do modelo de redes completamente aleatórias em descrevê-las, de fato, a melhor maneira de se testar uma hipótese é confrontá-la com o sistema natural objeto de estudo. Destacamos anteriormente a importância do grau médio $\langle k \rangle$ e como ele se relaciona com as distribuições Binomial e Poisson, sendo assim vamos tomar uma rede social mundial ($N = 7 \times 10^9$) como exemplo de rede real.

Sabendo que os sociólogos estimam que uma pessoa comum conhece mil pessoas pelo primeiro nome (BARABÁSI, 2013), vamos supor que $\langle k \rangle \approx 1000$. Se as conexões

dessa sociedade aleatória obedecerem a distribuição de Poisson podemos observar que os parâmetros estatísticos inerentes a distribuição proporcionam serias restrições as conexões da rede e conseqüentemente na maneira como as pessoas se socializam. Examinando a dispersão prevista pela distribuição para o caso hipotético proposto, se $\langle k \rangle \approx 1000$ os valores mínimo e máximo de k são respectivamente iguais a $k_{min} = 816$ e $k_{max} = 1185$, conforme apêndice B. Esses resultados indicam que o grau de um nó da rede deve residir em uma estreita faixa (816 a 1185), em outras palavras, indivíduos que chamam a atenção (outliers) por serem populares (muitos amigos) ou introvertidos (poucos amigos), simplesmente não existem nessa sociedade modelada por redes totalmente aleatórias. Todos possuem aproximadamente a mesma quantidade de conexões.

Essas constatações conflitam claramente com algumas redes reais já conhecidas, como o Facebook por exemplo, onde existem usuários com mais de cinco mil amizades (WILSON; GOSLING; GRAHAM, 2012). Uma breve inspeção no gráfico mostrado na Figura 10 demonstra que o modelo de redes totalmente aleatória não é adequado para modelar redes reais. Note que o ajuste de Poisson subestima o tamanho e a frequência dos nós com baixo e alto grau k e prevê um maior número de nós na vizinhança de $\langle k \rangle$.

Figura 10 – Distribuição de k para alguns casos listados na Tabela 1. A linha representa a distribuição de Poisson e os pontos dados experimentais, ambos com o mesmo $\langle k \rangle$.



Fonte: (BARABÁSI, 2013) - Adaptada

Na Tabela 1 apresentamos outros exemplos de redes reais, e todas apresentam o mesmo comportamento que sugere uma distribuição de k segundo uma lei de potência. Muito embora essas redes não sejam descritas satisfatoriamente pelo modelo de redes aleatórias, essas duas proposições podem possuir alguns pontos de similaridade ou até mesmo alguns parâmetros propositalmente escolhidos para servirem de referência. No

Tabela 1 – Redes reais e sua inadequação ao modelo de redes totalmente aleatórias.

Rede	N	L	$\langle k \rangle$	Referência
Internet	192.244	609.066	6,34	(CHEN et al., 2002)
Power Grid	4.941	6.594	2,67	(WATTS; STROGATZ, 1998)
Rede de Citação	783.339	6716198	8,57	(REDNER, 1998)
Interação Proteica	2.115	2.240	2,12	(JEONG et al., 2001)
Rede de Atores	449.913	25.516.482	113.43	(AMARAL et al., 2011)

próximo capítulo discutiremos esse assunto e como ele contribuiu para o surgimento de um novo modelo de rede conhecido como redes de mundo pequeno ou Small-World.

Fractais

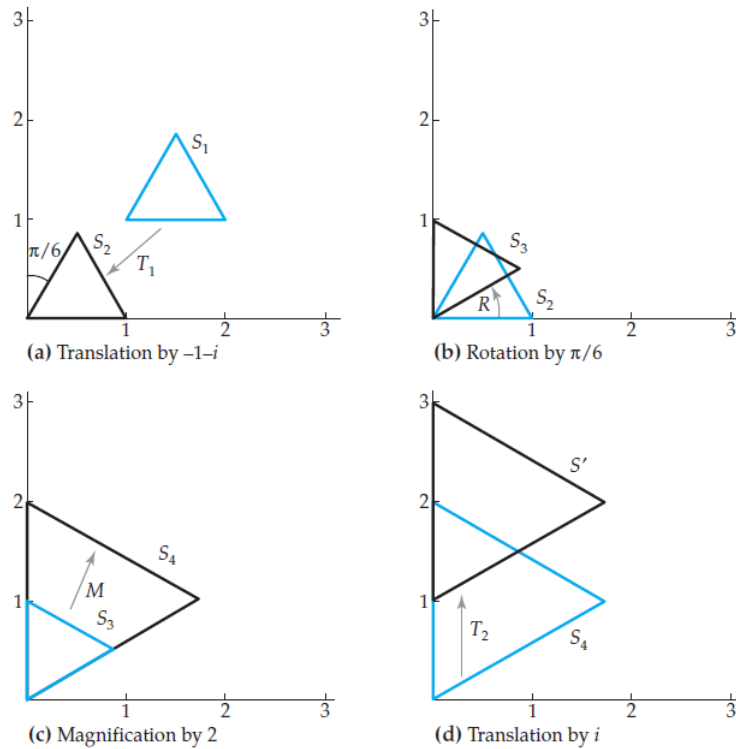
Fractais constituem uma classe de objetos geométricos complexos que frequentemente apresentam dimensão fracionária ([HAUSDORFF, 1918](#)). O termo fractal tem sua origem na palavra *fractus*, que no latim significa fragmento ou quebrado, e foi introduzido pela primeira vez por Mandelbrot ([MANDELBROT, 1977](#)). Antes de Mandelbrot os fractais já eram conhecidos, porém pouco estudados. O pioneirismo de seu trabalho possibilitou novas abordagens a sistemas físicos de difícil solução devido a enorme capacidade dos fractais em descrever objetos naturais com alta precisão, tornando-se ferramenta indispensável no estudo de sistemas complexos e desordenados ([FEDER, 2013](#); [MANDELBROT](#); [FRAME, 1987](#); [BUNDE](#); [HAVLIN, 2013](#); [BAISH](#); [JAIN, 2000](#); [TURCOTTE, 1989](#); [TURCOTTE, 1986](#); [NONNENMACHER](#); [LOSA](#); [WEIBEL, 2013](#); [GRIFFITH, 1994](#); [PIETRONERO, 1987](#); [PEEBLES, 1989](#); [JONES et al., 2005](#); [MEAKIN, 1998](#); [KENETT](#); [PERC](#); [BOCCALETTI, 2015](#)).

Neste capítulo, abordaremos as principais propriedades dos fractais, bem como a sua estrutura e algoritmo de construção. Seu comportamento frente as mudanças de escala é peça-chave para construção de mecanismos de busca eficiente que iremos propor nos capítulos subsequentes.

3.1 Auto-Similarity

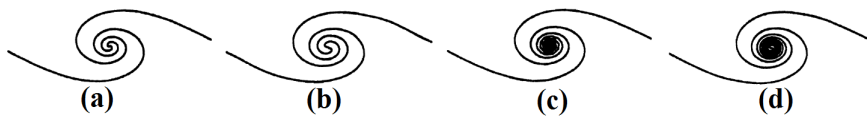
Na geometria Euclidiana elementar, as formas geométricas possuem similaridade (ou são semelhantes) quando apresentam a mesma forma. Especificamente, quando avaliamos dois objetos semelhantes, um pode ser obtido através do outro por meio da utilização de fatores de escala uniformemente distribuídos (aumento/redução), mesmo que ocorra rotação, translação ou reflexão especular conforme a [Figura 11](#).

Figura 11 – Transformações com a manutenção de similaridade no plano complexo: (a). Transformação de translação T_1 em $-1-i$. (b) Transformação de rotação R em $\pi/6$. (c) Transformação de escala M com aumento de $2x$. (d) Transformação de translação T_2 no eixo imaginário em i . Em todos os casos ilustrados, as estruturas, em azul e em preto, são semelhantes.



Fonte: (ZILL; SHANAHAN, 2009)

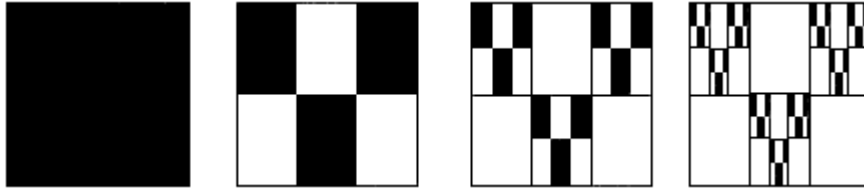
Figura 12 – Turbulência em fluidos é um exemplo de estruturas portadoras de auto-similaridade.



Fonte: (BIRDI, 1993)

Podemos ampliar a noção de similaridade ao domínio fractal, neste caso, o termo aplicado é *Auto-Similaridade*. Um objeto geométrico é dito portador de auto-similaridade quando apresentar invariância de forma ao ser submetido a uma transformação de escala, que é denominada de *Transformação de Similaridade*, onde o fator que conecta as estruturas original e transformada é denominado *Fator de Escala*. Na Figura 12 apresentamos um exemplo da ocorrência natural de estruturas auto-similares na turbulência de fluidos (CARR, 1997; DAUPHINÉ, 2013; ARMENTANO; LEGNANI; CYMBERKNOP, 2017).

Figura 13 – Auto-Afinidade para fractais anisotrópicos. A subdivisão do inicializador na horizontal utiliza um fator de escala 3, enquanto que na vertical o fator de escala utilizado é 2.



Fonte: (BEN-AVRAHAM; HAVLIN, 2000)

3.2 Auto-Afinidade

Quando pensamos nas *Transformações de Similaridade* de maneira mais abrangente, identificamos a necessidade de se considerar os casos onde o *Fator de Escala* não é uniformemente distribuído em todas as direções ao longo da estrutura do objeto. Em outras palavras, fractais com um certo grau de anisotropia podem apresentar invariância parcial de sua forma, a depender da transformação aplicada, manifestando uma simetria de dilatação anisotrópica. Nos referimos a esta característica como *Auto-Afinidade*. Dessa forma, as *Transformações de Auto-Similaridade* podem ser consideradas um caso especial das transformações de *Auto-Afinidade* onde os *Fatores de Escala* são os mesmos em todas as direções (CAMPBELL, 2014; YFANTIS; FLATMAN; ENGLUND, 1988).

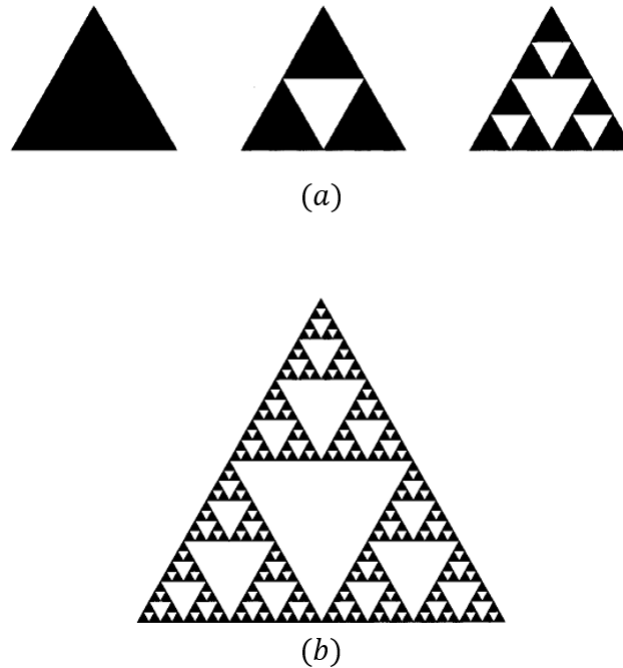
Conforme será discutido em detalhes na seção 3.3, um fractal pode ser construído através de um elemento fundamental, chamado de *Inicializador* que sofre infinitas transformações, podendo ser de auto-afinidade ou de auto-similaridade. Na Figura 13 apresentamos um exemplo utilizando um quadrado unitário como inicializador (PICKOVER, 1998; CRILLY; EARNSHAW; JONES, 2013). Note que são realizadas diferentes transformações de escala nas direções horizontal e vertical, onde foram aplicados fatores de escala iguais 3 e 2 respectivamente.

3.3 A construção de fractais

A definição precisa dos fractais, assim como todas as suas propriedades elementares demandam recursos matemáticos que estão muitos distantes dos propósitos desse trabalho, por isso optamos por explorar as principais funcionalidades desses objetos omitindo o formalismo matemático sempre que possível. Nesse sentido, assumimos que a criação de fractais requer apenas dois elementos: O *Inicializador* e o *Gerador*.

O *Inicializador* é uma estrutura euclidiana (1D, 2D ou 3D) tomada como elementar, que posteriormente sofrerá infinitas transformações de auto-similaridade ou auto-afinidade.

Figura 14 – Triângulo de Sierpinski: Formação do fractal com crescimento interior ao inicializador. (a) Inicializador e duas primeiras iterações. (b) Figura com melhor aproximação do fractal verdadeiro, registrada na quinta iteração.

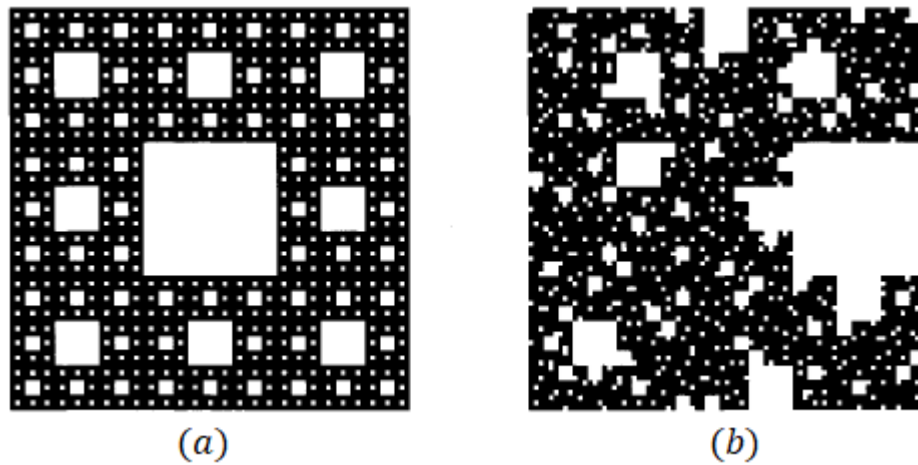


Fonte: ([BEN-AVRAHAM; HAVLIN, 2000](#))

O *Gerador*, por outro lado, é o algoritmo que descreve precisamente a transformação aplicada. Comumente, os geradores dispõem de regras fixas a cada iteração ou são dotados de componentes aleatórios, isso dá origem aos fractais denominados de *Determinísticos* e *Estatísticos*, respectivamente. Na [Figura 14](#) apresentamos um exemplo de criação de um fractal determinístico bastante conhecido, denominado *Triângulo de Sierpinski* ([GOODCHILD, 1980](#); [CARPINTERI; MAINARDI, 2014](#)). O inicializador é um triângulo unitário, e o gerador é o algoritmo que funciona identificando o ponto médio de todas as arestas, conectando-os para criar outros 4 triângulos menores para finalmente remover o triângulo central. Note que as ilustrações apresentadas na figura ainda não são o fractal em questão, tratando-se apenas de meras aproximações que melhoram a cada iteração, ou seja, fractais são figuras geométricas que só existem no limite quando o número de iterações tende ao infinito. Isso tem implicações profundas na implementação computacional de simulações que são modeladas por fractais, uma vez que não dispomos de memória infinita para utilização de fractais no sentido literal ([VOSS, 1991](#); [CHARKALUK; BIGERELLE; IOST, 1998](#)). Dessa forma, nos contentamos em utilizar estruturas aproximadamente fractal e avaliamos seu comportamento limite, conforme será mostrado em seções futuras.

Outro caso importante de fractal determinístico é conhecido como *Quadrado de Sierpinski*, que inclusive é objeto de nossa pesquisa. Nele podemos constatar que o

Figura 15 – Fractais determinísticos e estatísticos para o quadrado de Sierpinski.



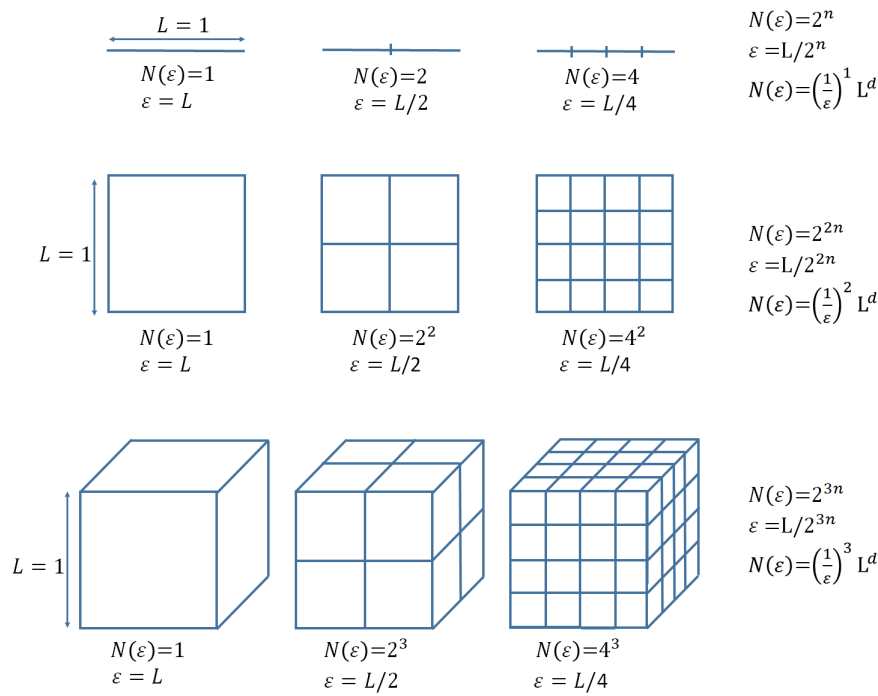
Fonte: (BEN-AVRAHAM; HAVLIN, 2000)

inicializador é um quadrado unitário, e o gerador consiste em dividir o quadrado original em 9 quadrados menores e iguais onde finalmente o quadrado central é removido (LAM; SANDER, 1992; MATSUSHITA; MEAKIN, 1988). A estrutura determinística resultante desse algoritmo se encontra na parte (a) da Figura 15. Note que se ao invés de remover sistematicamente apenas o quadrado central a cada iteração, removêssemos qualquer um dos nove quadrados escolhidos aleatoriamente produziríamos o fractal estatístico que consta na parte (b) da figura. É evidente que com a o aumento das iterações a quantidade de buracos na estrutura cresce consideravelmente, o que diminui exponencialmente a sua densidade, conforme será visto nas próximas seções. Este parâmetro deve ser avaliado com muito cuidado durante a implementação da navegação em redes com geometria fractal, pois o aumento excessivo do tamanho da rede e/ou nível de aproximação do fractal verdadeiro pode elevar o número de restrições (buracos) a níveis proibitivamente elevados (LIU, 1986).

3.4 A Dimensão Fractal

Uns dos principais descritores de um fractal é a sua dimensão d_f (VICSEK, 1992). Como discutido anteriormente, a medição do volume de fractais imersos no espaço euclidiano de dimensão d nos permite concluir que fractais são objetos de dimensão fracionária. Atualmente a literatura oferece diversas formas de obtenção da dimensão fractal, utilizaremos aqui a mesma abordagem proposta em (HAUSDORFF, 1918). Avaliamos inicialmente objetos geométricos da geometria euclidiana e catalogamos suas relações de dimensionalidade, conforme mostrado na Figura 16. De uma forma geral, tomamos um objeto unitário no qual realizamos diversas transformações de auto-similaridade por um fator ε , depois contabilizamos o número de estruturas semelhantes $N(\varepsilon)$.

Figura 16 – Relações de dimensionalidade.



Fonte: Autor

É possível verificar que a relação entre ε e $N(\varepsilon)$ é dependente da dimensão do elemento inicializador d , e que de forma geral podemos representar tal relação da seguinte forma:

$$N(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^d L^d. \quad (3.1)$$

Assumindo que fractais manifestam a mesma propriedade, temos que

$$N(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^{d_f} L^{d_f}. \quad (3.2)$$

a dimensão fractal de Hausdorff é dada através da manipulação da [Equação 3.2](#)

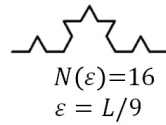
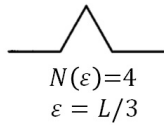
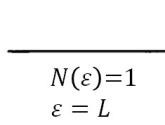
$$d_f = \frac{\log N(\varepsilon)}{\log L + \log(1/\varepsilon)} \quad (3.3)$$

Considerando que L é desprezível para pequenos valores de ε , a dimensão fractal se reduz a

$$d_f = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}. \quad (3.4)$$

Figura 17 – Dimensão fractal.

Curva de Koch

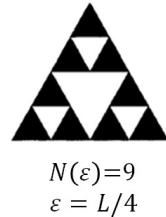
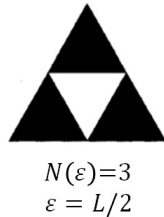
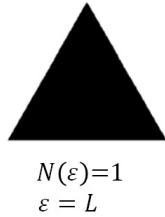


$$N(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^{d_f} L^{d_f}$$

$$4 = \left(\frac{1}{L/3}\right)^{d_f} L^{d_f}$$

$$d_f = \frac{\log 4}{\log 3} = 1,26186 \dots$$

Triângulo de Sierpinski

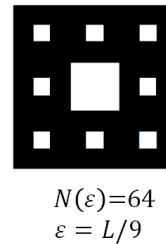
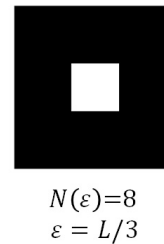
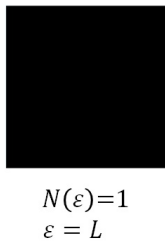


$$N(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^{d_f} L^{d_f}$$

$$3 = \left(\frac{1}{L/2}\right)^{d_f} L^{d_f}$$

$$d_f = \frac{\log 3}{\log 2} = 1,58496 \dots$$

Quadrado de Sierpinski



$$N(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^{d_f} L^{d_f}$$

$$8 = \left(\frac{1}{L/3}\right)^{d_f} L^{d_f}$$

$$d_f = \frac{\log 8}{\log 3} = 1,89279 \dots$$

Fonte: Autor

Na [Figura 17](#) apresentamos alguns exemplos de aplicação da [Equação 3.4](#) nos principais fractais já catalogados, onde destaca-se o *Quadrado de Sierpinski*, um vez que nossas simulações serão realizadas em redes com essa geometria.

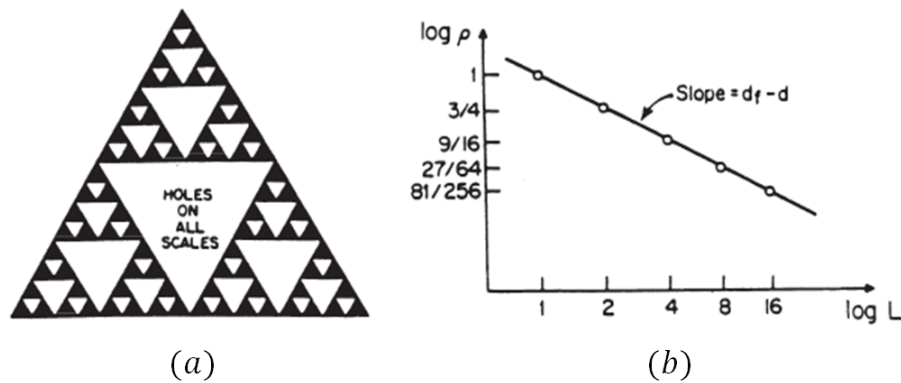
3.5 A densidade de Fractais

Conforme mencionamos antes, a densidade dos fractais é um descritor de extrema relevância. Curiosamente, ao contrário do que se observa na geometria euclidiana, a densidade de objetos fractais varia com a escala. Para exemplificar essa importante propriedade vamos examinar o Triângulo de Sierpinski. A definição operacional de densidade aplicada a fractais é apresentada na [Equação 3.5](#)

$$\rho(L) \equiv \frac{M(L)}{V(L)}. \quad (3.5)$$

Uma vez que o inicializador é um objeto projetado no plano bidimensional euclidiano a densidade se torna $\rho(L) \equiv \frac{M(L)}{L^2}$.

Figura 18 – A densidade em fractais: (a) Estrutura repleta de buracos após a quarta iteração. (b) Decaimento da densidade a cada iteração. Plotagem dos dados da Tabela 2.



Fonte: (BUNDE; HAVLIN, 2013)

Como podemos observar na Figura 18, a cada iteração ocorre um incremento exponencial de buracos no objeto, o que faz com que sua densidade decaia violentamente. Destacamos que essa característica, inerente a todos os fractais, deve ser monitorada com bastante critério. Nos capítulos seguintes iremos mostrar que a navegação em redes com geometria fractal utilizam esses buracos para restringir o estabelecimento de conexões de longo alcance e que isso eleva o custo computacional da operação. Vale lembrar que na prática nunca trabalhamos com fractais exatos, uma vez que estes são estruturas que só existem no limite de infinitas iterações. Quanto maior o número de iterações melhor será essa aproximação, no entanto, é preciso haver equilíbrio em relação a complexidade do algoritmo. Em nossa abordagem, será mostrado que a navegação continua exigindo redes de grande porte, mas que não é necessário extrapolar o limite computacional em busca de um número excessivo de iterações.

Tabela 2 – Decaimento da densidade fractal para o Triângulo de Sierpinski.

Iteração	Comprimento L	Massa	Densidade $\rho(L)$
0	2^0	3^0	$(\frac{3}{4})^0$
1	2^1	3^1	$(\frac{3}{4})^1$
2	2^2	3^2	$(\frac{3}{4})^2$
3	2^3	3^3	$(\frac{3}{4})^3$

Fonte: Produzido pelo autor.

Algoritmos de geração de redes artificiais

"Everything should be made as simple as possible, but no simpler."

Albert Einstein

Neste capítulo discutiremos alguns algoritmos para geração de redes artificiais, será mantida a mesma metodologia de apresentação aplicada em capítulos anteriores. Partiremos de um modelo básico e rudimentar em termos de representatividade das redes reais, com correção gradual das principais inconsistências através de novos modelos. O leitor deve se atentar ao largo espectro de descobertas e avanços na área de ciência de redes acumulados até a atualidade, e que neste trabalho abordaremos apenas a revisão bibliográfica essencial para nossa pesquisa.

4.1 Redes de mundo pequeno (Small-World)

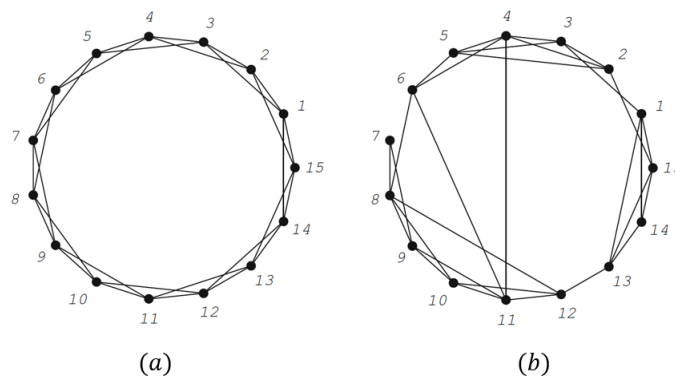
No capítulo anterior, vimos que o modelo de ER não era adequado para representar algumas redes reais. Foi mostrado que as redes de colaboração científica e de conexão entre atores, por exemplo, portam a curiosa propriedade de possuir pequeno diâmetro e alto coeficiente de agrupamento. Qualquer rede de grande porte que reúna essas propriedades é denominada rede de mundo pequeno ou small-world. De fato, esse fenômeno já era amplamente conhecido como *seis graus de separação* pelo público em geral e um dos primeiros estudos sociológicos publicados a seu respeito foi conduzido por Milgram (TRAVERS; MILGRAM, 1967). Esse artigo será de grande importância para a evolução dos modelos que vamos propor neste trabalho.

Tabela 3 – Métricas utilizadas por Watts e Strogatz para identificar e simular redes de mundo pequeno artificialmente.

Tipo de rede	$\langle L \rangle$	$\langle C \rangle$
Aleatória (ER)	Pequeno	Pequeno
Regular	Grande	Grande
Small-World	Pequeno	Grande

Watts e Strogatz foram os primeiros a construir um algoritmo capaz de gerar artificialmente redes com essas características (WATTS; STROGATZ, 1998). Através das informações constantes na Tabela 3 eles identificaram que a aleatoriedade estava diretamente ligada a $\langle L \rangle$ mais curtos. Sabendo que o aumento da aleatoriedade em redes regulares reduz consideravelmente o valor de $\langle L \rangle$ mantendo as demais características (BOLLOBÁS; CHUNG, 1988), eles propuseram um processo de construção de redes que tem início com uma rede totalmente regular onde as ligações são refeitas de forma aleatória (criação de atalhos), o que como já se sabe, reduz apenas o valor de $\langle L \rangle$, tornando a rede original numa rede de pequeno mundo. Por simplicidade, as simulações são realizadas em redes unidimensionais com aplicação de condições de contorno periódicas (WATTS, 2004), ou seja, os vértices estão alinhados sobre uma linha onde as duas extremidades são unidas. Isso confere a forma de anel ou de ringue a rede, de modo que cada nó é conectado a $2m$ vizinhos mais próximos, m a direita e m a esquerda.

Figura 19 – Método de Watts e Strogatz para geração de redes de mundo pequeno com $N = 15$ e $m = 2$: (a) Rede totalmente regular com $p = 0$ (b) Rede com algumas ligações refeitas com probabilidade $p = 0,2$.

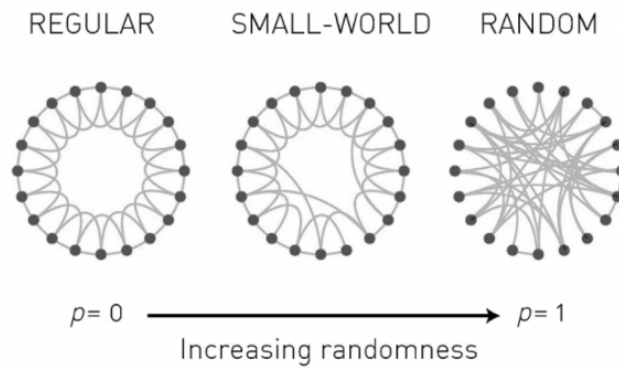


Fonte: (BOCCARA, 2010) - Adaptada

O algoritmo segue percorrendo o anel e em cada vértice uma ligação é refeita com probabilidade p ou preservada com $1 - p$. Autoconexões não são permitidas e nem múltiplas conexões entre o mesmo par de vértices, a Figura 19 ilustra um exemplo. De modo geral, o modelo de Watts e Strogatz é uma interpolação entre os modelos de redes regular e aleatória de acordo com a Figura 20. Já a Figura 21 mostra os resultados obtidos

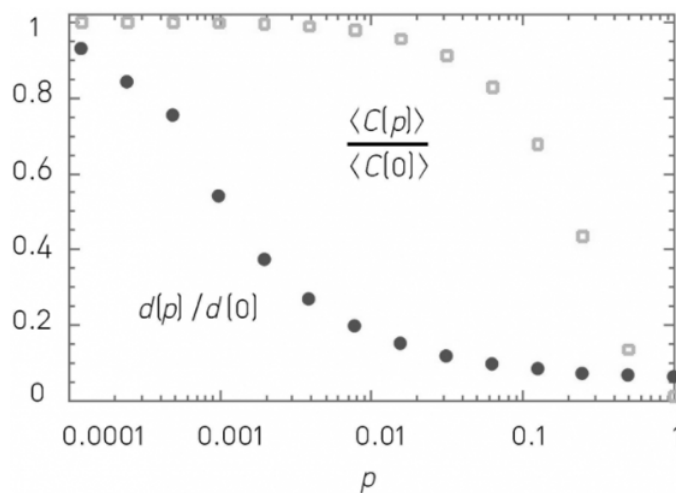
por eles em simulações de computador.

Figura 20 – Abordagem de Watts e Strogatz para gerar redes de mundo pequeno através de interpolação de redes com links refeitos com probabilidade p .



Fonte: (BARABÁSI, 2013) - Adaptada

Figura 21 – Resultados de simulações produzidas com o modelo de Watts e Strogatz para redes de mundo pequeno.



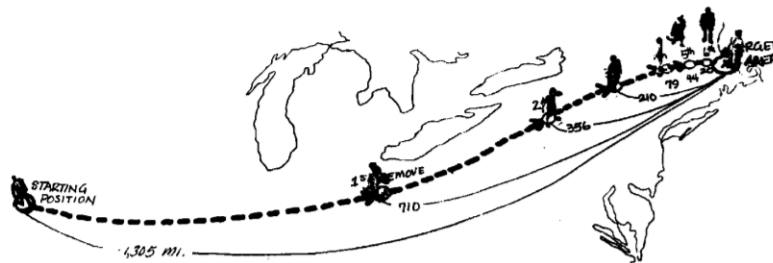
Fonte: (BARABÁSI, 2013) - Adaptada

O modelo de Watts e Strogatz foi uma contribuição inestimável para desenvolvimento da ciência de rede, no entanto está longe de ser considerado um modelo geral e totalmente acabado. Ainda existem lacunas a serem preenchidas. Os atalhos inseridos justificam a melhoria da navegabilidade devido ao encurtamento do trajeto, ou seja, sempre existirá um caminho mais curto para se trafegar dentro da rede, e é justamente sobre isso que se trata o trabalho de Milgram. Misteriosamente o caminho mais curto é mais provável de ser encontrado, mesmo o navegante detendo apenas informações locais sobre a rede. Esse fato só é explicado por outro modelo (Kleinberg) que será abordado na [seção 4.3](#).

4.2 O experimento de Milgram

O experimento de Milgram ([TRAVERS; MILGRAM, 1967](#)) foi um estudo empírico conduzido nos Estados Unidos nos anos de 1960 com o intuito de investigar o fenômeno do mundo pequeno, ou *seis graus de separação*. Utilizou-se do envio de mensagens através de cartas como metodologia para se investigar o número de intermediários entre o remetente primário e o destinatário final, os portadores da mensagem só poderiam passá-la a diante através de um de seus contatos e o pouco conhecimento que eles detinham eram os dados do destinatário. A alta taxa de insucesso no recebimento das cartas pelo destinatário, cerca de 97%, colocou a crença na existência dos *seis graus de separação* em posição de fragilidade. Outros pesquisadores e grandes corporações como o Facebook e o próprio Watts tentaram recriar o experimento através de ferramentas modernas e outras metodologias mas também não obtiveram sucesso e o problema segue em aberto.

Figura 22 – Experimento de Milgram: A distância euclidiana entre o portador da mensagem e o alvo diminui a metade conforme a mensagem se aproxima do alvo.



Fonte: ([TRAVERS; MILGRAM, 1967](#))

Mesmo assim, o trabalho de Milgram foi de fundamental importância no entendimento e desenvolvimento do algoritmo de geração de redes de pequeno mundo. A [Figura 22](#) mostra um comportamento peculiar na trajetória da mensagem, e nela vemos que inicialmente o portador da mensagem está a 1305 milhas do alvo, no passo seguinte está a 710 e depois a 356 e segue diminuindo por um fator de aproximadamente dois até atingir o alvo. Esse fato evidenciou que os portadores da mensagem deveriam estar levando em consideração a distância geográfica para escolher dentre seus contatos a melhor opção de envio.

4.3 Redes de Kleinberg

O modelo de Kleinberg tem como principal objetivo simular a navegação em redes de pequeno mundo em complemento a proposta de Watts e Strogatz discutida em seções anteriores. Em seu livro (WATTS, 2004), Watts discute seu próprio modelo (WATTS; STROGATZ, 1998) e avalia que o próximo passo evolutivo reside na forma com a qual os atalhos são gerados, uma vez que não há garantia alguma de que os caminhos mais curtos serão encontrados, mesmo que eles existam em abundância. De fato, o trabalho de Milgram (TRAVERS; MILGRAM, 1967) evidenciou que a utilização de uma aleatoriedade uniforme na criação de atalhos não leva em consideração os efeitos da noção de distância física que o navegante possui ao transitar através da rede. De modo geral, sistemas complexos são comumente modelados por grafos cuja conectividade desempenha um papel mais abrangente, a noção de distância deixa de ser restrita ao espaço físico para se tornar mais abstrata e genérica com o intuito de representar uma ampla variedade de sistemas, como por exemplo: Classes sociais, profissões, interações moleculares e assim por diante. Tudo isso deixa claro que o modelo vai muito além do campo da sociologia.

Estas evidências levaram Kleinberg (KLEINBERG, 1999; KLEINBERG, 2000) a propor uma construção teórica baseada no modelo de redes de mundo pequeno de Watts e Strogatz (1998) visando modelar a dinâmica de rede incluindo atalhos aleatórios, distâncias geográficas e as condições necessárias para que uma rede seja navegável.

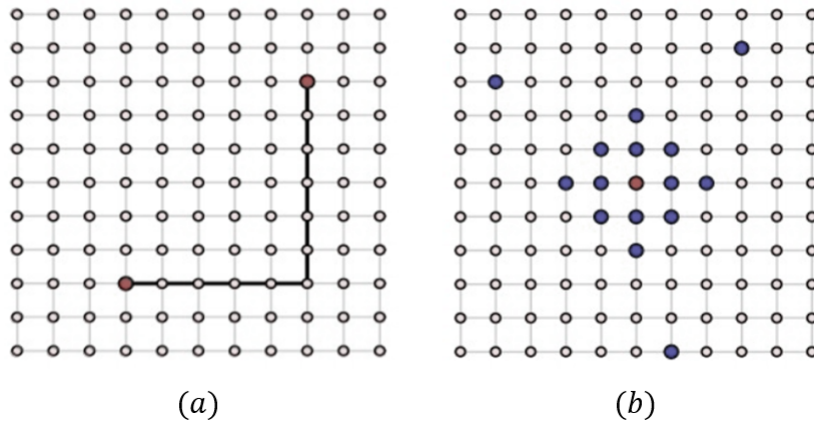
4.3.1 Construção do modelo

Kleinberg deu início a construção de seu modelo delimitando a vizinhança do portador da mensagem (contatos), ele assumiu que qualquer nó possui conexão com todos os nós no interior de uma região, cuja distância escolhida é p . Em seguida delimitou o número q de links de longo alcance (atalhos) que podem existir em cada vértice e estabeleceu que eles ocorreriam segundo uma distribuição de probabilidade que privilegia os nós mais próximos ao portador da mensagem, conforme Equação 4.1,

$$P(u, v) \sim \frac{1}{d(u, v)^\alpha}. \quad (4.1)$$

Por exemplo, na parte (a) da Figura 23 a distância entre dois nós genéricos u e v , denotada por $d(u, v)$ é igual a 11, já na parte (b) o portador da mensagem aparece de vermelho e seus contatos de azul. Sua vizinhança tem no máximo uma distância de duas unidades ($p = 2$) e lhe é permitido a criação de três atalhos ($q = 3$) escolhidos aleatoriamente segundo a distribuição apresentada na Equação 4.1. Note que ainda estamos sob as regras do modelo de redes de *Mundo Pequeno*, já que a princípio, as conexões de longo alcance (atalhos) continuam presentes e mantém o seu caráter aleatório. A contribuição

Figura 23 – Determinação dos contatos do portador da mensagem: (a) A distância entre dois nós u e v , destacados de vermelho, é $d(u, v) = 11$. (b) O ponto vermelho é o portador da mensagem e em azul seus contatos, sendo $p = 2$ e $q = 3$.



Fonte: Autor

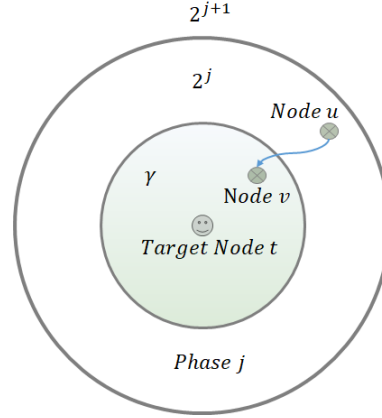
de Kleinberg para generalização do modelo de Watts (WATTS; STROGATZ, 1998) tem como ponto central a inclusão de uma distribuição de probabilidade $P(u, v)$ que regula a criação de atalhos entre os nós u e v . Quanto maior a distância $d(u, v)$ entre os pontos u e v , menos provável será a inclusão de um atalho entre eles. Na Equação 4.1, $\alpha \geq 0$ é um parâmetro denominado expoente de clusterização, que pode ser ajustado para tornar a rede navegável, ou ainda, demonstrar a generalidade da abordagem de Kleinberg. Fazendo $\alpha = 0$ a probabilidade de criação de atalhos é constante e uniformemente distribuída, reduzindo a rede ao modelo de Watts e Strogatz (mundo pequeno). Em outro extremo, se tomarmos $\alpha \rightarrow \infty$, a probabilidade de um link de longo alcance ser estabelecido tende a zero, o que nos levaria a uma rede regular. Tal abordagem, para certos valores de α , temos que os atalhos existem em abundância mas são difíceis de se encontrar, para outros, os atalhos não podem ser encontrados por que eles simplesmente não existem. De certa forma, a navegação parece ser regida por "foças" conflitantes que atingem o equilíbrio quando os caminhos curtos são fáceis de encontrar. Redes que atingem esta condição especial são ditas *redes navegáveis*.

4.3.2 Tempo de Entrega

O número de intermediários computados durante o envio das cartas no estudo de Milgram é denominado *tempo de entrega* T . Como visto, a navegação em redes de mundo pequeno está intimamente ligada ao algoritmo de construção da rede, a inclusão de atalhos acrescenta caminhos curtos e o ponto central da pesquisa de Kleinberg foi estabelecer quais seriam as condições necessárias para minimizar o tempo de entrega, mesmo com o conhecimento limitado e local dos intermediários. Algoritmos com tais características são

denominados de descentralizados ou gulosos.

Figura 24 – A navegação de Kleinberg estratifica a trajetória da mensagem em LnN fases das quais podem ocorrer até LnN passos, no pior caso.



Fonte: Autor

O exame criterioso da [Equação 4.1](#) e de seus desdobramentos permite uma compreensão mais ampla dos mecanismos envolvidos na navegação. Sendo $P_{(u,v)} = 1/Ar_{(u,v)}^\alpha$, onde $A = \sum_{u \neq v} r_{(u,v)}^{-\alpha}$ é uma constante de normalização que pode ser obtida de forma aproximada no limite do contínuo em que d é a dimensão da rede e L seu comprimento ¹:

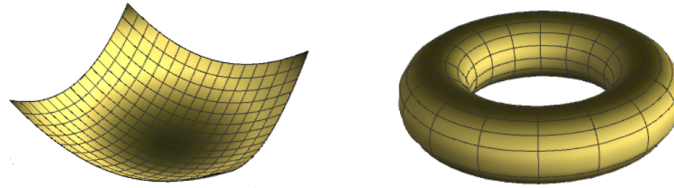
$$A \sim \int_1^L r^{-\alpha} r^{d-1} dr \sim \begin{cases} L^{d-\alpha}, & \alpha < d, \\ \ln L, & \alpha = d, \\ (\alpha - d)^{-1}, & \alpha > d. \end{cases} \quad (4.2)$$

Por simplicidade Kleinberg adotou em suas simulações $d = 2$, $p = q = 1$ e o modelo toroidal de rede mostrado na [Figura 25](#). Ele também visou alguma consistência com as observações de Milgram projetando a busca descentralizada de modo a estratificar a trajetória do navegante em regiões geográficas chamadas de *fase*, como mostra a [Figura 24](#).

Assumindo que cada fase é uma região circular e que todas as fases atribuídas a uma trajetória são organizadas de forma concêntrica, ele construiu um sistema semelhante ao que é ilustrado na [Figura 24](#). Isso possibilitou a execução de cálculos estatísticos que demonstraram analiticamente que a navegação é estratificada em LnN fases, podendo ocorrer LnN passos em cada uma delas no pior caso. Fornecemos uma prova para esses resultados no [Apêndice C](#). Note que a última fase a ser navegada ocorre quando o intermediário v e o nó alvo t se encontrarem na mesma fase, que vamos denominar aqui como fase γ . As conclusões das simulações computacionais que resultaram desse estudo

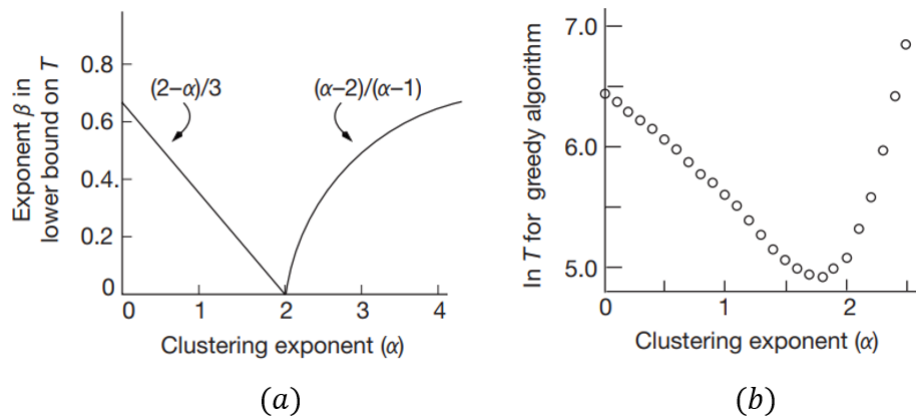
¹ Note que N ou L podem ser utilizados indiscriminadamente nas expressões que se seguirão, uma vez que $L \sim N^{1/d}$

Figura 25 – Condições de contorno periódicas aplicadas em métodos tradicionais de simulação. Também conhecida como modelo toroidal de rede.



Fonte: Autor

Figura 26 – Navegação de Kleinberg em redes de pequeno mundo com $d = 2$, $p = 1$ e $q = 1$. Prova computacional de que a rede se torna navegável em $\alpha = 2$.



Fonte: (KLEINBERG, 2000)

teórico se encontram resumidas na Figura 26. Nas partes (a) e (b) da figura, vemos que o tempo de entrega é minimizado quando $\alpha = 2$, caso especial onde $d = 2$. Notoriamente antes e depois de d o tempo de entrega é maior porque os atalhos possuem relativa dificuldade de serem encontrados, o que torna a rede *pesquisável* apenas quando $\alpha = d$. Por causa da Equação 4.2, quando a rede não é pesquisável ($\alpha \neq d$) o tempo de entrega deixa de ser minimizado e passa a ser proporcional a uma potência de base N da forma $T \sim N^\beta \sim L^\beta$. De fato, a rede é considerada pesquisável (ou navegável) quando $\alpha = d = 2$ e toda a análise correlata ao tempo de entrega pode ser sintetizada pelas expressões presentes na Equação 4.3.

$$\begin{aligned}
 T &\sim L^{(2-\alpha)/3}, & \text{case } 0 \leq \alpha < 2, \\
 T &\sim \ln^2 L, & \text{case } \alpha = 2, \\
 T &\sim L^{(\alpha-2)/(\alpha-1)}, & \text{case } 2 < \alpha < 3.
 \end{aligned}
 \tag{4.3}$$

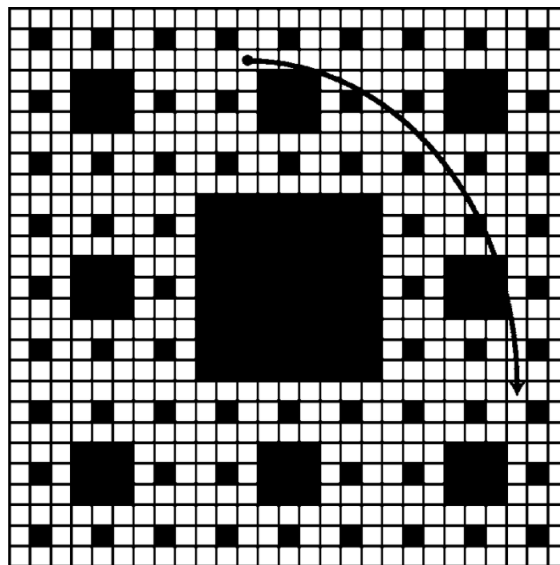
Note que esse estudo pode ser generalizado para geometrias diferentes da euclidiana

(BARTHELEMY, 2018), para os nossos propósitos de pesquisa vamos expandir nossa abordagem para a geometria fractal de modo que $d = d_f$, onde d_f é a dimensão fractal.

4.4 Navegação em redes de mundo pequeno fractal

Ainda utilizando o modelo toroidal da Figura 25, pesquisadores (ROBERSON; BEN-AVRAHAM, 2006) conseguiram generalizar o modelo de Kleinberg para uma rede fractal conhecida como quadrado de sierpinski. Na Figura 27 mostramos a ilustração utilizada pelos autores em seu artigo original, nela podemos verificar a criação de um link de longo alcance e que nem todos os sítios são navegáveis (pintados de preto).

Figura 27 – Criação de link de longo alcance entre sítios do quadrado de sierpinski



Fonte: (ROBERSON; BEN-AVRAHAM, 2006)

É importante notar que fractais são figuras impossíveis de serem representadas em sua totalidade uma vez que elas existem no limite $L \rightarrow \infty$ e seria necessário memória também infinita para executar as simulações, além disso, também seria inviável a aplicação do modelo toroidal da Figura 25. No entanto Roberson e Ben-Avraham propuseram uma técnica chamada *finite-size corrections* que permite trabalhar com versões reduzidas e aproximadas do fractal, como consequência a dimensão d em que o roteamento ocorre não é a dimensão fractal d_f , mas sim um valor aproximado que melhora conforme a rede simulada se aproxima do fractal verdadeiro, ou seja, $d \rightarrow d_f$ quando $L \rightarrow \infty$. Baseados na Equação 4.4 eles propuseram que α_{min} do fractal exato é uma função de N tal que $\alpha_{min}(N) \sim \alpha(\infty) + A/(\ln N)^2$ onde $\alpha(\infty) = \lim_{N \rightarrow \infty} \alpha_{min}(N) = d_f$. Essa abordagem permite encontrar d_f com um simples ajuste linear de curva e será utilizada para validar

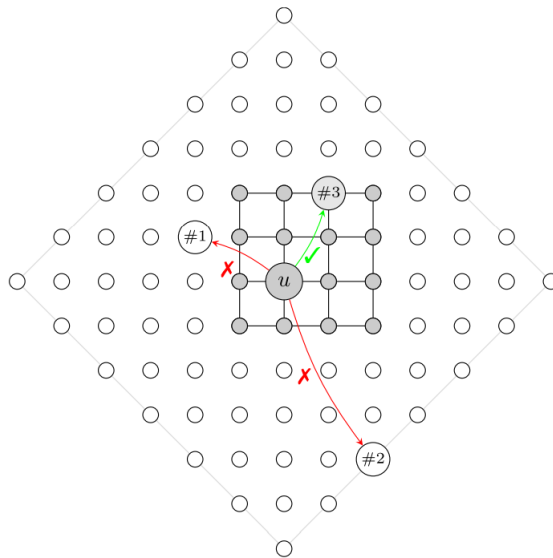
nossas simulações. Também foi demonstrado que as equações [Equação 4.2](#) e [Equação 4.3](#) são generalizações para redes fractais, basta substituir $d = d_f$. Assim obtemos a [Equação 4.4](#)

$$\begin{aligned} T &\sim L^{(d_f-\alpha)/(1+d_f)}, & \text{case } 0 \leq \alpha < d_{df}, \\ T &\sim \ln^2 L, & \text{case } \alpha = d_{df}, \\ T &\sim L^{(\alpha-d_f)/(\alpha-d_f+1)}, & \text{case } d_{df} < \alpha < 3. \end{aligned} \quad (4.4)$$

4.5 Amostragem por rejeição dinâmica

A amostragem por rejeição dinâmica é um algoritmo desenvolvido por ([MATHIEU; COMTE, 2020](#)) para desenhar atalhos em simulações de roteamento em redes de pequeno mundo e é conhecida por apresentar boa complexidade computacional. É importante reforçar que nessa nova abordagem o modelo de Kleinberg é mantido intacto no sentido de que não existe nenhum tipo de aproximação numérica. A diferença entre as duas abordagens fica restrita a maneira com a qual a rede é modelada: Kleinberg utiliza o modelo toroidal da [Figura 25](#), enquanto que Mathieu e Comte utilizam *máscaras de aceitação*. Será mostrado que a última abordagem não deforma a rede, ao contrário da primeira, e também conta com a utilização de um pequeno banco de dados reutilizável que armazena todos os cálculos de probabilidade de criação de atalhos, o que dispensa repetições desnecessárias.

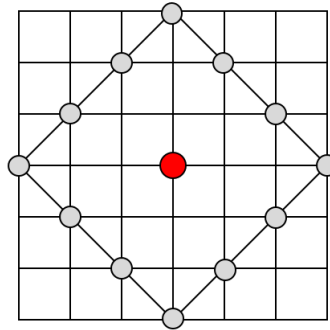
Figura 28 – Máscara de aceitação com $L = 4$: Ao centro está a rede G representada na cor cinza e sobreposta a rede B_u (sítios pintados de branco). Vértices fora do domínio de G não estão aptos a receberem a mensagem.



Fonte: ([MATHIEU; COMTE, 2020](#))

As redes que iremos tratar nesse trabalho serão denotadas por $G(L, \alpha, p, q)$. Como já

Figura 29 – Para cada vértice existem outros $4i$ vértices vizinhos a uma distância i . A vizinhança do nó de cor vermelha, separada por três arestas $i = 3$ é 12.



Fonte: Autor

definimos antes, L é o comprimento da rede, α o expoente de clusterização, p o comprimento máximo da vizinhança do nó corrente e q o número de links de longo alcance permitidos. A novidade aqui é que G apresentará um comportamento dinâmico e será sobreposta a uma segunda rede virtual e estática B_u cujos pontos estão dentro de uma bola de raio $2(n - 1)$ capaz de incluir todos os pontos de G . O centro de B_u é tomado como centro relativo e a partir dele são desenhados todos os $2(n - 1)$ atalhos com probabilidades proporcionais a $d(u, v)^{-\alpha}$. Além disso, os autores idealizaram uma metodologia de sorteio para a escolha dos próximos portadores da mensagem baseados no desacoplamento entre as componentes radial e angular de suas posições com o intuito de reduzir a complexidade do algoritmo de $O(n)$ para $\Omega(n^2)$. Sendo assim, para cada raio i existem $4i$ nós (Figura 29), que serão amostrados uniformemente através de uma variável angular. Durante a execução do algoritmo, G se desloca sobre B_u de modo que o portador da mensagem $u \in G$ sempre ocupe o centro de B_u . Os atalhos de B_u , criados anteriormente, podem ser reaproveitados desde que pertençam a G . Por isso, dizemos que a rede G age como uma máscara de aceitação nesse processo. Na Figura 28 apresentamos o exemplo utilizado pelos autores, nele $L = 4$ e as duas primeiras tentativas de criação de atalho são frustradas porque os candidatos se encontravam fora do domínio de G .

Parte II

Proposições e Simulações

Máscaras de aceitação fractal

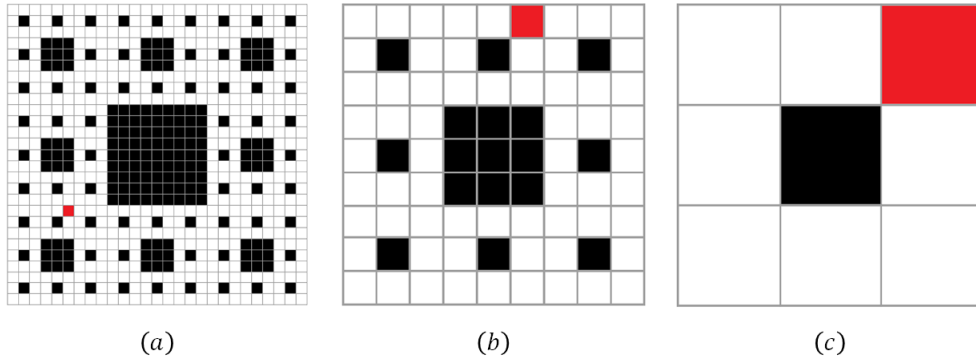
5.1 Introdução

Para gerar atalhos em redes com topografia fractal utilizando a *Amostragem por Rejeição Dinâmica* discutida em capítulos anteriores, é necessário que as máscaras de aceitação reconheçam a topografia fractal com o intuito de detectar pontos específicos, que neste trabalho, vamos chamar de *pontos fractais*. Por inspeção, podemos afirmar que a detecção desses pontos é fortemente dependente do tipo de fractal selecionado para estudo. Este capítulo dará ênfase ao fractal denominado como *quadrado de sierpinski*, cujos pontos fractais se encontram pintados de preto na [Figura 30](#). Também oferecerá detalhes da construção das máscaras de aceitação aplicáveis a redes desse tipo.

5.2 Algoritmo de busca fractal (Fractal Search)

Como visto na [seção 4.5](#), máscaras de aceitação para redes bidimensionais introduzem restrição a criação de links de longo alcance. Em nossas simulações estendemos essa restrição aos pontos fractais. Além disso, publicações anteriores ([ROBERSON; BEN-AVRAHAM, 2006](#)) demonstram que a navegação em redes fractais demandam tamanhos extremamente elevados e conseqüentemente uma enorme quantidade de sítios ($L \approx 1,2 \times 10^9$, $1,4 \times 10^{17}$). Também é importante notar que redes dessa magnitude possuem o número de pontos fractais bastante elevado devido ao decaimento da densidade discutido na [seção 3.5](#), em outras palavras, existem muitas restrições ao estabelecimento efetivo de um link de longo alcance. Para incorporar os pontos fractais às máscaras de aceitação, considerando todos esses aspectos operacionais, desenvolvemos uma rotina booleana de busca capaz de localizar todos os pontos fractais da rede. Visando a otimização do uso de memória,

Figura 30 – Aplicação de *Fractal_Search*: Redução da geração do fractal $n = 3, 2, 1$ nas pastes (a), (b) e (c) respectivamente. Como o ponto v pintado de vermelho não se encontra na região central de nenhuma das redes G_3, G_2, G_1 a rotina informa que v é navegável.



Fonte: Autor

a possibilidade da criação de um banco de dados para armazenar na íntegra todas as informações da rede foi afastada, ao invés disso preferimos uma metodologia de busca dinâmica aplicável a cada nó intermediário candidato a receber o atalho. O algoritmo foi projetado para utilizar a autossimilaridade intrínseca do *Quadrado de Sierpinski* aliada propriedades recursivas com o intuito de minimizar o custo computacional da busca. Ele conta com a função $Fractal_Search(x, y, L_n)$ que recebe as coordenadas (x, y) do ponto $v \in G_n$ de comprimento L_n , $n = 1, 2, 3, \dots$ onde n é a geração do fractal segundo a sua lei de formação. $Fractal_Search(x, y, L_n)$ retorna *True* sempre que v for um ponto fractal e *False* caso contrário.

Ainda sobre a lei de formação, a rede aproximadamente fractal G_n precisa ter um comprimento L_n governado por uma exponencial de base 3, tal que $L_n = 3^n$. Dessa forma, $G_n \rightarrow G$ quando $n \rightarrow \infty$, em outras palavras, G_n são as redes finitas e simuláveis que iremos utilizar e elas se aproximam do fractal exato (infinito) quando a geração n for grande o suficiente para fazer surgir as propriedades previstas por Kleinberg. Conforme mencionamos antes, em nossa construção, os sítios pintados de preto (pontos fractais) não podem receber links de longo alcance. A varredura tem início com a rede em seu comprimento original $L_n = 3^n$ e a cada looping recursivo G_n é dividido em 9 partes de tamanhos $L_n/3$ onde a geração n é diminuída em 1 unidade. Note que a região central de todas as redes $G_n, G_{(n-1)}, G_{(n-2)}, \dots, G_1$ nunca está apta a receber atalhos e *Fractal_Search* testa primeiramente se v pertence a essa região, em caso negativo ela segue sua marcha recursiva diminuindo a geração do fractal até o limite de $n = 1$, onde a testagem final é realizada e a rotina finalmente informa se v é um ponto fractal ou não. Na [Figura 30](#) oferecemos um exemplo de aplicação, em uma pequena rede G_3 de comprimento $L_3 = 3^3$, queremos que *Fractal_Search* informe se o nó v (pintado de vermelho) é um ponto fractal.

Na parte (a), a busca tem início com G_3 , como v não ocupa a parte central de G_3 a rotina segue para a próxima etapa recursiva de menor geração extraindo uma cópia da nona parte de G_3 que contém v , conforme mostra a parte (b). O processo se repete até $n = 1$, ilustrada na parte (c), onde se conclui que v não ocupa a região central e portanto não se trata de um ponto fractal. No final desta seção fornecemos ao leitor um pseudocódigo do algoritmo *Fractal Search*.

Algoritmo 1: Fractal Search

Input: x , y and L

Output: True or False

Function fractal_search(x , y , L_n):

$L_{(n-1)} \leftarrow L_n/3$

$x_id \leftarrow x//L_{(n-1)}$

$y_id \leftarrow y//L_{(n-1)}$

if $x_id == 1$ **and** $y_id == 1$ **then**

 | **return** True

else if $L_{n-1} \leq 1$ **then**

 | **return** False

else

 | $x_new \leftarrow x \% L_{(n-1)}$

 | $y_new \leftarrow y \% L_{(n-1)}$

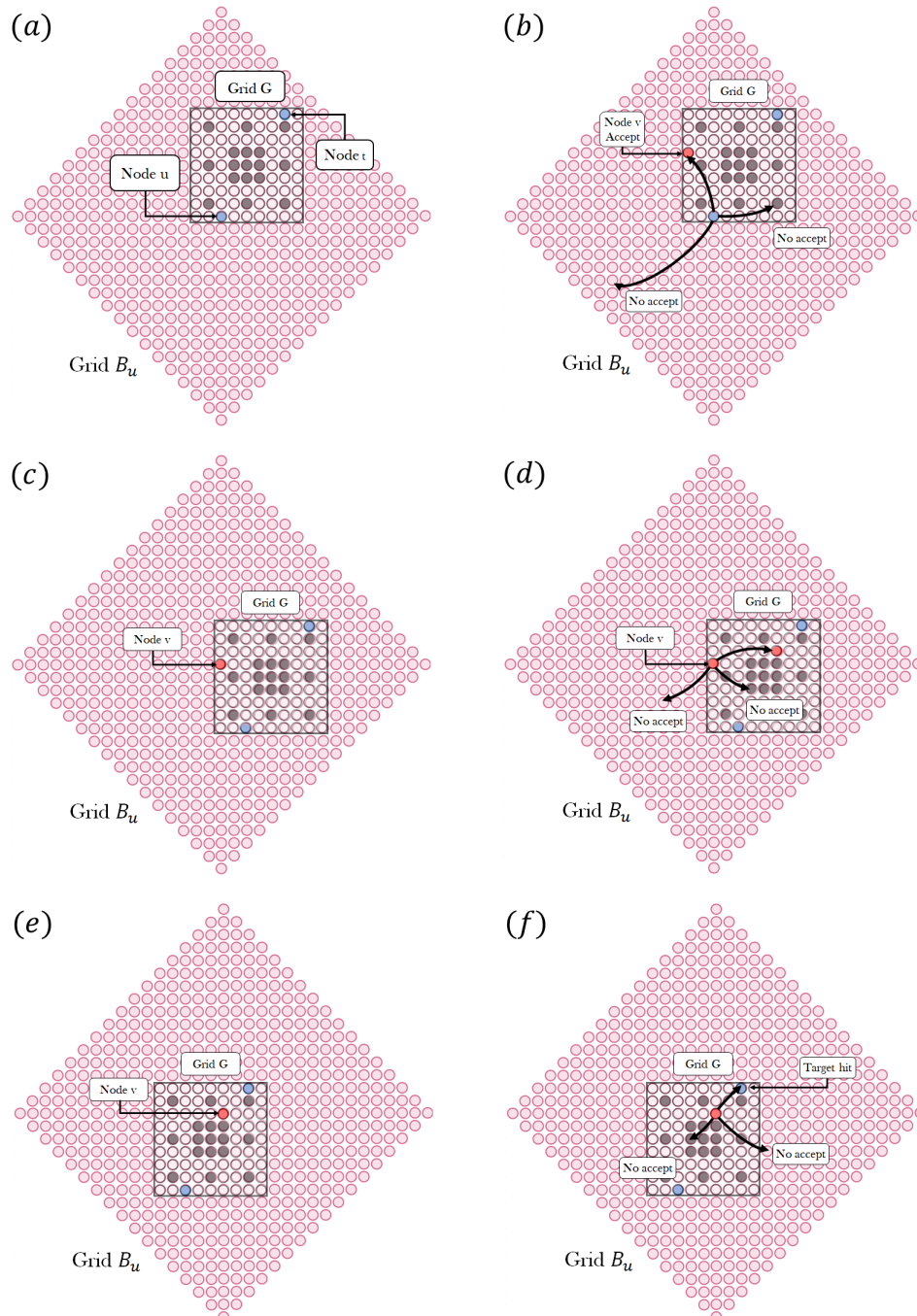
endif

return fractal_search(x_new , y_new , $L_{(n-1)}$)

5.3 Máscaras de aceitação fractal

Máscaras de aceitação, como o próprio nome sugere, oferece restrições a escolhas de nós candidatos a receberem atalhos. Nesta seção vamos propor uma generalização no modelo original de Mathieu e Comte para que as máscaras de aceitação reconheçam a geometria fractal do *Quadrado de Sierpinski* e imponham as restrições adicionais pertinentes. O resultado disso fica melhor explicado através do exemplo ilustrado na [Figura 31](#). A rede G de comprimento L é nosso objeto de simulação, adequando a notação claramente temos G_3 e $L_2 = 3^2$ respectivamente. A rede virtual B_u aparece pintada na cor rosa, note que o impedimento para o estabelecimento de atalhos entre u e $v \notin G$ permanece inalterado em relação ao método tradicional. No entanto, os vértice $v \in G$ pintados de cinza, são reconhecidos como pontos fractais e não poderão receber atalhos, utilizamos *Fractal Search* para identificar esses sítios. Os demais vértices $v \in G$ permanecem navegáveis e aptos a receber atalhos como originalmente seria. A busca descentralizada segue esse padrão de seleção e navegação até atingir o alvo. É importante ressaltar que os pontos fractais são navegáveis, porém não através de links de longo alcance.

Figura 31 – Amostragem por rejeição dinâmica em redes com geometria fractal: O portador da mensagem $u \in G$ sempre ocupa a posição central de B_u e reutiliza alguns de seus atalhos já calculados anteriormente e de maneira única. Na abordagem tradicional apenas os nós $v \in G$ estão aptos a reutilizar esses atalhos e isso não muda para a abordagem fractal. No entanto, nem todos os pontos de G são navegáveis devido a restrições da própria geometria do fractal. Nesta figura ilustramos a dinâmica de criação de atalhos em uma rede $L = 3^2$.



Fonte: Autor

Resultados numéricos e discussões

6.1 Parâmetros adotados na simulação e regime computacional

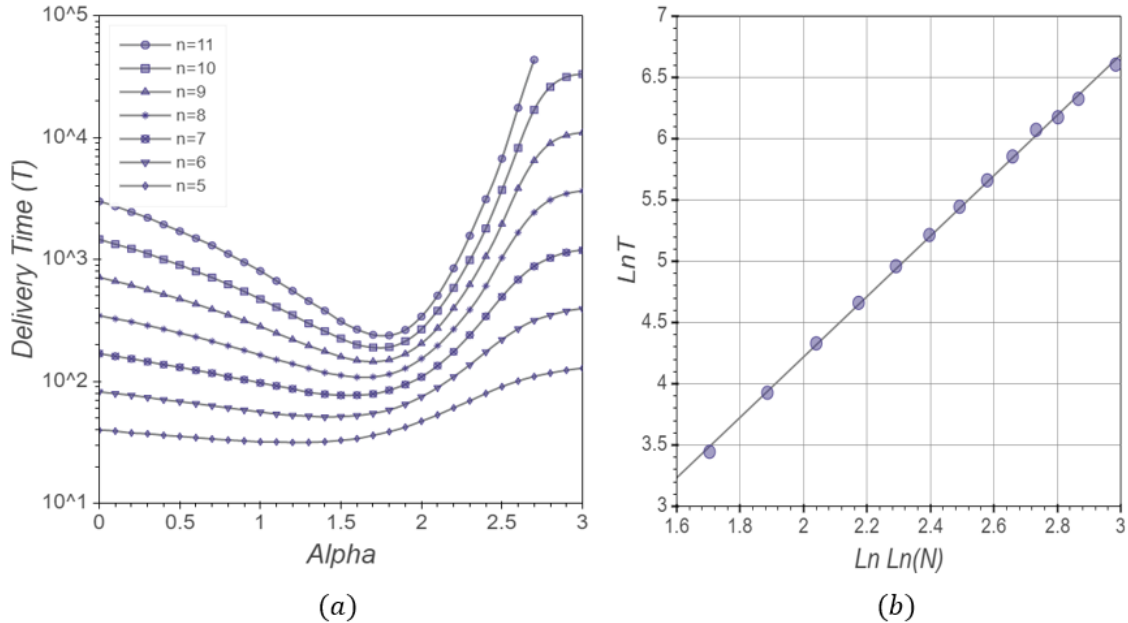
Iniciamos a nossa pesquisa simulando G_n para algumas gerações n do quadrado de sierpinski, tais que $n \in \Theta = [5, 11]$ e calculamos o tempo de entrega tomando a média de 1000 execuções para cada um dos valores de $\alpha \in \Gamma = [0, 3]$ num passo de 0,02, em todos os casos trabalhamos com $p = q = 1$. Utilizamos computador desktop com processador Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz RAM 16.0 GB. Todas as simulações foram realizadas na linguagem Python3.

6.2 Resultados numéricos

6.2.1 A existência de um α que minimiza T e consistência com o modelo de Kleinberg

O primeiro passo em nossa investigação computacional foi simular a navegação de Kleinberg em redes de mundo pequeno com geometria fractal utilizando a amostragem por rejeição dinâmica acoplada a *Fractal Search* (AMARAL; BELICH, 2021), os resultados se encontram na [Figura 32](#). Na parte (a) verifica-se a existência de um valor de α que minimiza o tempo de entrega, vamos nos referir a estes valores como α_{min} . Note que para todos os valores de n adotados ($L = 3^n$) verifica-se que $\alpha_{min} \neq d_f = 1,89278$, mas isso não implica em violação das previsões teóricas de Kleinberg, pelo contrário, todas essas observações eram esperadas e a justificativa para esse comportamento reside no fato de estarmos trabalhando com aproximações simuláveis do fractal exato ($n \rightarrow \infty$),

Figura 32 – Navegação de Kleinberg em redes de mundo pequeno com geometria fractal utilizando a amostragem por rejeição dinâmica acoplada a *Fractal Search*: (a) Tempo de entrega para redes $L = 3^n$ tal que $n \in \Theta$: Em todos os casos $\alpha_{min} \neq d_f$ sem violar o roteamento de Kleinberg. Comportamento sugere que α_{min} converge para d_f com o aumento de n . (b) Consistência da simulação com a previsão $T \sim (LnN)^2$ apresentada na [Equação 4.4](#).



Fonte: Autor

conforme discutido no [Capítulo 3](#). A parte (b) trata da consistência de nossas simulações com a [Equação 4.4](#), ou seja, quando a rede é navegável temos que $T \sim (lnN)^2$. Então, tomando o logaritmo em ambos os lados da equação (ou utilizando eixos em escala log-log) espera-se que a plotagem desses parâmetros forneça um gráfico em forma de reta conforme o obtido. Além disso, uma breve inspeção na parte (a) da figura, sugere que α_{min} se aproxima de d_f mediante ao aumento de n . Na verdade, essa é uma propriedade já catalogada em navegações por redes fractais e era esperada, uma vez que o aumento de n melhora a qualidade da aproximação fractal. Na próxima seção apresentaremos estudos mais aprofundados sobre essa convergência com o intuito de verificar seu comportamento diante das alterações propostas pelas máscaras de aceitação.

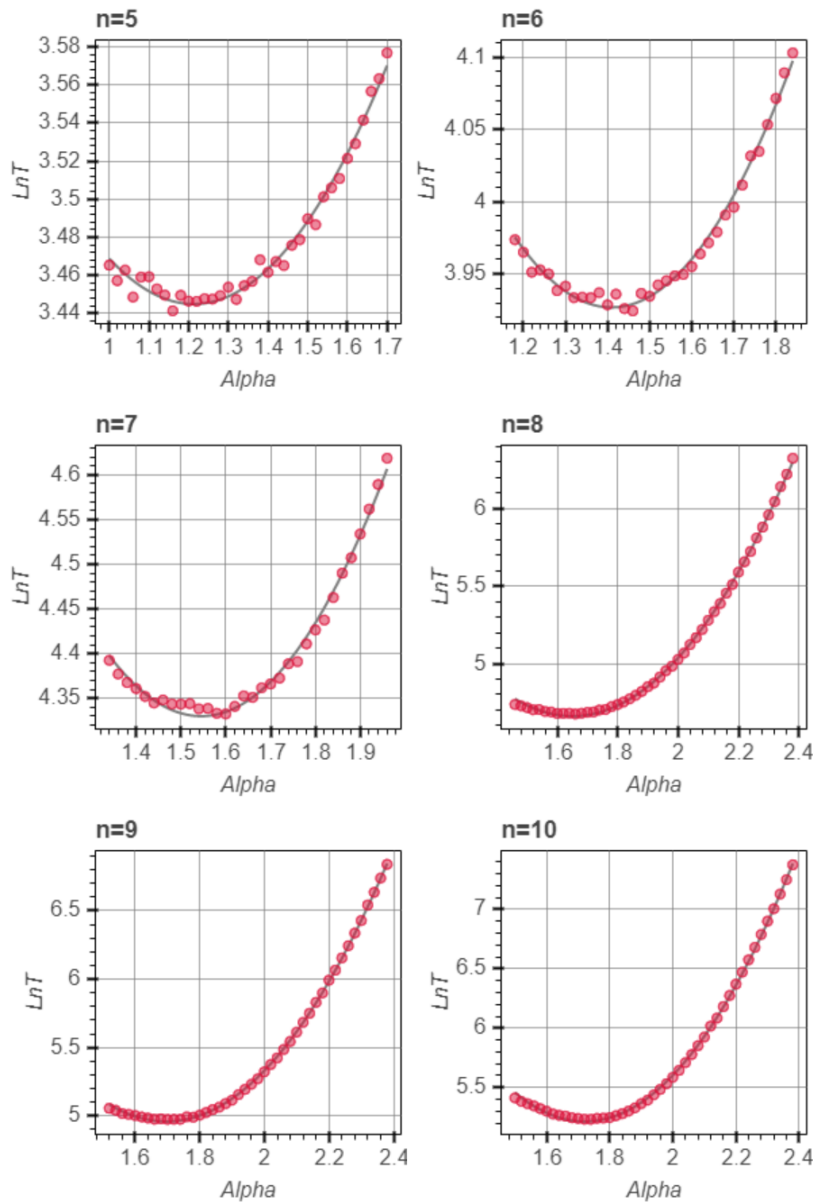
6.2.2 A convergência $d \rightarrow d_f$

Nesta subseção vamos aplicar um método denominado *finite-size corrections* (ROBERSON; BEN-AVRAHAM, 2006), discutido na [seção 4.4](#), para verificar se convergência $d \rightarrow d_f$ detectada nos resultados apresentados na subseção anterior realmente indicam convergência. Iniciamos essa tarefa determinando os valores de α_{min} para cada um dos

valores de n simulados, conforme consta na [Figura 33](#) e em nossa publicação ([AMARAL; BELICH, 2021](#)). Isso foi feito através de um ajuste de curva parabólica considerando valores de α nas proximidades de α_{min} com $n \in [5, 16]$. Na [Figura 33](#) mostramos alguns dos ambientes de plotagem para esses ajustes. Para $5 \leq n \leq 7$ os ajustes são aceitáveis se tornando ótimos para $n \geq 8$, adotando o número de realizações fixo em $R = 1000$.

Utilizamos esses dados na aplicação do método *finite-size corrections* e verificamos que realmente a convergência existe e que ela conduz ao valor correto da dimensão fractal. Na [Figura 34](#) obtemos a dimensão fractal com um desvio percentual de 0,095%. Na [Figura 35](#), utilizamos uma escala logarítmica no eixo das abscisas para facilitar a visualização da convergência, isso deixa claro que a utilização de redes com tamanhos excessivamente elevados não são vantajosas porque despendem um enorme esforço computacional em troca de ganhos ínfimos na melhoria da aproximação.

Figura 33 – Navegação de Kleinberg em redes de mundo pequeno com geometria fractal utilizando a amostragem por rejeição dinâmica acoplada a *Fractal Search*: Ajuste parabólico de curva para obtenção de α_{min} .

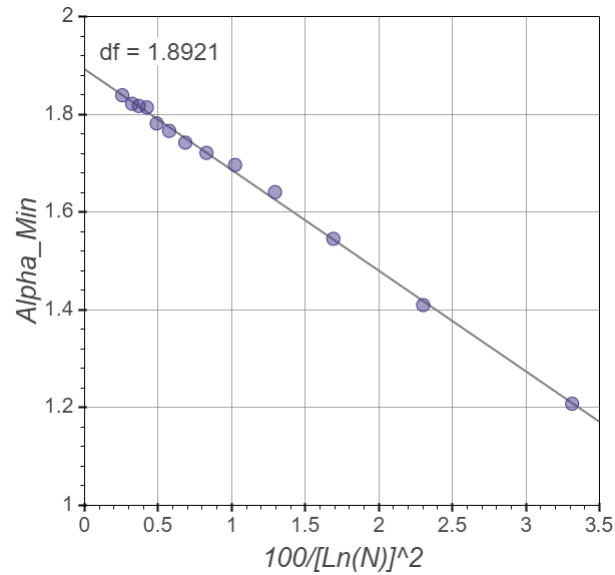


Fonte: Autor

6.2.3 Análise de desempenho

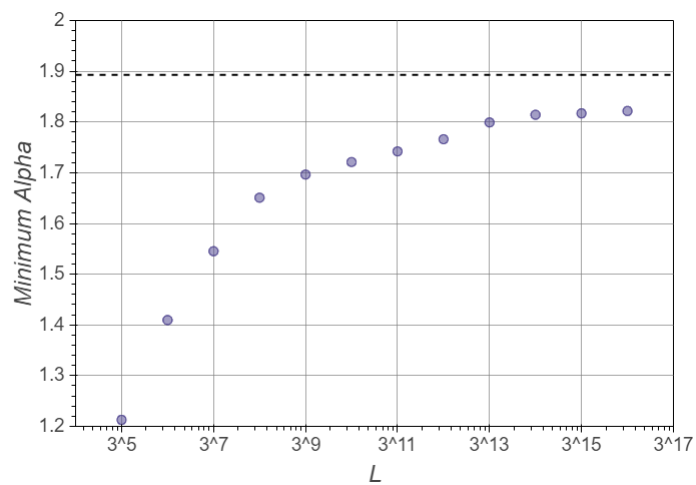
Como visto, para $R = 1000$, simulações como $n = 5, 6$ produzem resultados aceitáveis. No entanto, para critérios mais conservadores a utilização de $n \geq 8$ é recomendada (AMARAL; BELICH, 2021). Evidentemente, o aumento no número de realizações R tem o potencial de melhorar a qualidade do ajuste mostrado na Figura 34 as custas do aumento do tempo de execução. Note ainda que a escolha do tamanho da rede através da geração n , deve ser feita com bastante critério. A utilização de tamanhos excessivamente grandes deixa de ser vantajosa devido a baixa contribuição para a convergência, conforme mostra a Figura 35. Além disso, o método *finite-size corrections* torna essa prática completamente desaconselhável. Realizamos testes de desempenho entre as técnicas convencionais, utilizando condições de contorno periódicas, e a que abordamos em nossa proposta de pesquisa com máscaras de aceitação acopladas a *Fractal Search*. O resultado se encontra na Figura 36 e podemos observar que nossa proposta apresentou desempenho superior para todas as configurações simuladas.

Figura 34 – Obtenção da dimensão fractal d_f pelo método *Finite-Size Corrections*.



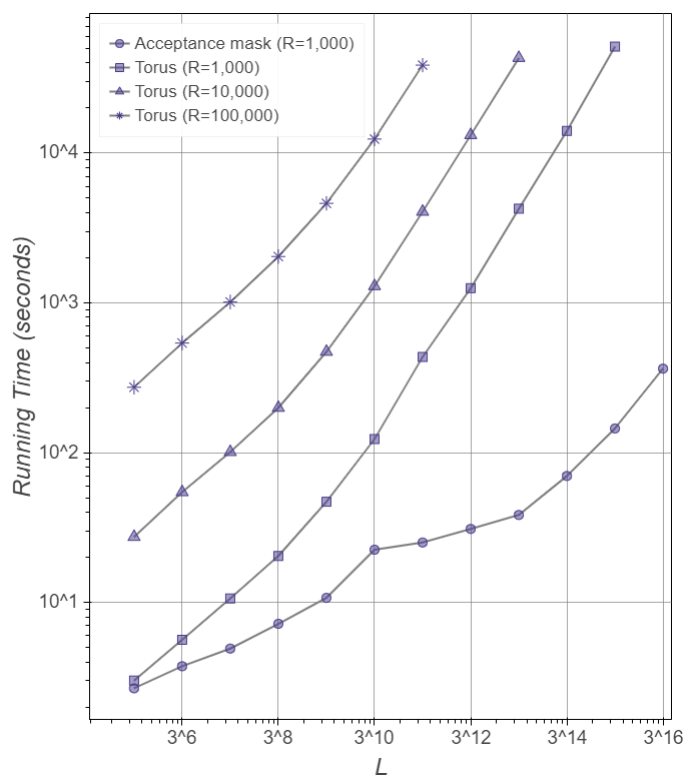
Fonte: Autor

Figura 35 – A convergência $d \rightarrow d_f$ é lenta. Tornando o trabalho com redes de comprimento L excessivamente grandes contra-indicado.



Fonte: Autor

Figura 36 – Análise do tempo de execução.



Fonte: Autor

Conclusão

Em nossa pesquisa, descobrimos que a simulação em redes de mundo pequeno com geometria fractal do tapete de Sierpinski, utilizando amostragem de rejeição dinâmica com máscaras de aceitação acopladas ao algoritmo *Fractal Search*, mostrou-se uma alternativa viável e mais eficiente do que as técnicas tradicionais. Ao longo do texto, ficou demonstrado que redes fractais devem ser de grande porte para que os fenômenos previstos por Kleinberg possam emergir e que isso pode impactar negativamente no tempo de execução de algoritmos de busca devido ao decaimento da densidade fractal. Propomos um algoritmo que contorna essas dificuldades (*Fractal Search*) e verificamos que ele obteve êxito em estender a abordagem de Mathieu e Comte para redes fractais. Por fim, utilizamos as principais previsões teóricas e computacionais já publicadas em periódicos para validar nossos resultados e verificamos que a abordagem proposta foi bem-sucedida. Além disso, identificamos que os critérios de aceitação relacionados a qualidade da simulação em computador deve levar em consideração o tempo de execução, o número de realizações R e o comprimento L da rede. Em todos os casos considerados, as máscaras de aceitação fractal apresentaram desempenho consideravelmente superior a abordagem tradicional onde se aplica as condições de contorno periódicas (torus). Chamamos a atenção para a enorme gama de aplicações físicas dessa técnica, tais como: teoria da percolação, fenômenos críticos e criticalidade, bem como problemas típicos de sistemas complexos e desordenados. Trabalhos futuros poderão ser desenvolvidos nessas áreas.

Referências

- AMARAL, L. et al. Classes of small-world networks. In: *The Structure and Dynamics of Networks*. [S.l.]: Princeton University Press, 2011. p. 207–210. Citado na página 45.
- AMARAL, L. A.; BELICH, H. Kleinberg’s navigation in fractal small-world networks by dynamic rejection sampling. *Brazilian Journal of Physics*, 2021. ISSN 1678-4448. Disponível em: <<https://doi.org/10.1007/s13538-021-00995-4>>. Citado 3 vezes nas páginas 73, 75 e 77.
- ARMENTANO, R. L.; LEGNANI, W.; CYMBERKNOP, L. J. Fractal analysis of cardiovascular signals empowering the bioengineering knowledge. *Fractal Anal-Appl Health Sci Soc Sci*, 2017. Citado na página 48.
- BAISH, J. W.; JAIN, R. K. Fractals and cancer. *Cancer research*, AACR, v. 60, n. 14, p. 3683–3688, 2000. Citado na página 47.
- BARABÁSI, A.-L. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society Publishing, v. 371, n. 1987, p. 20120375, 2013. Citado 9 vezes nas páginas 29, 35, 36, 40, 42, 43, 44, 57 e 95.
- BARNES, J. A.; HARARY, F. Graph theory in network analysis. *Social networks*, Elsevier, v. 5, n. 2, p. 235–244, 1983. Citado na página 32.
- BARRAT, A.; BARTHELEMY, M.; VESPIGNANI, A. *Dynamical processes on complex networks*. [S.l.]: Cambridge university press, 2008. Citado na página 31.
- BARTHELEMY, M. *Morphogenesis of spatial networks*. [S.l.]: Springer, 2018. Citado na página 63.
- BEN-AVRAHAM, D.; HAVLIN, S. *Diffusion and reactions in fractals and disordered systems*. [S.l.]: Cambridge University Press, 2000. Citado 3 vezes nas páginas 49, 50 e 51.
- BIRDI, K. Fractals in chemistry, geochemistry, and biophysics. 1993. Citado na página 48.
- BOCCARA, N. *Modeling complex systems*. [S.l.]: Springer Science & Business Media, 2010. Citado na página 56.
- BOLLOBÁS, B. *Modern graph theory*. [S.l.]: Springer Science & Business Media, 2013. v. 184. Citado na página 29.
- BOLLOBÁS, B.; CHUNG, F. R. K. The diameter of a cycle plus a random matching. *SIAM Journal on discrete mathematics*, SIAM, v. 1, n. 3, p. 328–333, 1988. Citado na página 56.

- BOLLOBÁS, B.; RIORDAN, O. M. Mathematical results on scale-free random graphs. *Handbook of graphs and networks: from the genome to the internet*, Wiley-Vch Weinheim, Germany, p. 1–34, 2003. Citado na página 35.
- BONACICH, P. Power and centrality: A family of measures. *American journal of sociology*, University of Chicago Press, v. 92, n. 5, p. 1170–1182, 1987. Citado na página 37.
- BORGATTI, S. P.; EVERETT, M. G. A graph-theoretic perspective on centrality. *Social networks*, Elsevier, v. 28, n. 4, p. 466–484, 2006. Citado na página 36.
- BULA, W.; NIKIEL, J.; TYMCHATYN, E. The königsberg bridge problem for peano continua. *Canadian Journal of Mathematics*, Cambridge University Press, v. 46, n. 6, p. 1175–1187, 1994. Citado na página 28.
- BUNDE, A.; HAVLIN, S. *Fractals in science*. [S.l.]: Springer, 2013. Citado 2 vezes nas páginas 47 e 54.
- CAMPBELL, S. C. *Mysterious patterns: finding fractals in nature*. [S.l.]: Boyds Mills Press, 2014. Citado na página 49.
- CARPINTERI, A.; MAINARDI, F. *Fractals and fractional calculus in continuum mechanics*. [S.l.]: Springer, 2014. v. 378. Citado na página 50.
- CARR, J. R. Statistical self-affinity, fractal dimension, and geologic interpretation. *Engineering geology*, Elsevier, v. 48, n. 3-4, p. 269–282, 1997. Citado na página 48.
- CHARKALUK, E.; BIGERELLE, M.; IOST, A. Fractals and fracture. *Engineering Fracture Mechanics*, Elsevier, v. 61, n. 1, p. 119–139, 1998. Citado na página 50.
- CHEN, Q. et al. The origin of power laws in internet topologies revisited. In: IEEE. *Proceedings. twenty-first annual joint conference of the ieee computer and communications societies*. [S.l.], 2002. v. 2, p. 608–617. Citado na página 45.
- COHEN, R.; HAVLIN, S. *Complex networks: structure, robustness and function*. [S.l.]: Cambridge university press, 2010. Citado na página 28.
- CRILLY, A. J.; EARNSHAW, R.; JONES, H. *Applications of fractals and chaos: the shape of things*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 49.
- DAUPHINÉ, A. *Fractal geography*. [S.l.]: John Wiley & Sons, 2013. Citado na página 48.
- ERDÖS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, v. 6, p. 290–297, 1959. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, v. 6, p. 290–297, 1959. Citado na página 23.
- ERDÖS, P.; RÉNYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, v. 5, p. 17–61, 1960. Citado na página 23.
- ERDÖS, P.; RÉNYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, v. 5, n. 1, p. 17–60, 1960. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, v. 38, p. 343–347, 1961. Citado na página 40.

- ERDÖS, P.; RÉNYI, A. On the strength of connectedness of a random graph. *Citeseer*, 1961. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. Asymmetric graphs. *Acta Mathematica Acad. Sci. Hungarica*, Springer, v. 14, p. 295–315, 1963. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. On random matrices. *Publ. Math. Inst. Hung. Acad. Sci.*, n. 8, p. 455–461, 1966. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungary*, v. 17, p. 359–368, 1966. Citado na página 40.
- ERDÖS, P.; RÉNYI, A. On random matrices ii. *Studia Sci. Math. Hungar.*, v. 13, p. 459–464, 1968. Citado na página 40.
- ESTRADA, E.; BODIN, Ö. Using network centrality measures to manage landscape connectivity. *Ecological Applications*, Wiley Online Library, v. 18, n. 7, p. 1810–1825, 2008. Citado na página 36.
- EULER, L. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, p. 128–140, 1741. Citado na página 29.
- FEDER, J. *Fractals*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 47.
- FOULDS, L. R. *Graph theory applications*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 30.
- FREEMAN, L. C. A set of measures of centrality based on betweenness. *Sociometry*, JSTOR, p. 35–41, 1977. Citado na página 37.
- GERHARDT, G. J.; LEMKE, N.; CORSO, G. Network clustering coefficient approach to dna sequence analysis. *Chaos, Solitons & Fractals*, Elsevier, v. 28, n. 4, p. 1037–1045, 2006. Citado na página 33.
- GIBBONS, A. *Algorithmic graph theory*. [S.l.]: Cambridge university press, 1985. Citado na página 32.
- GILBERT, E. N. Random graphs. *The Annals of Mathematical Statistics*, JSTOR, v. 30, n. 4, p. 1141–1144, 1959. Citado 2 vezes nas páginas 23 e 39.
- GONDRAN, M.; MINOUX, M.; VAJDA, S. *Graphs and algorithms*. [S.l.]: John Wiley & Sons, Inc., 1984. Citado na página 29.
- GOODCHILD, M. F. Fractals and the accuracy of geographical measures. *Journal of the International Association for Mathematical Geology*, Springer, v. 12, n. 2, p. 85–98, 1980. Citado na página 50.
- GRIBKOVSKAIA, I.; SR, Ø. H.; LAPORTE, G. The bridges of königsberg—a historical perspective. *Networks: An International Journal*, Wiley Online Library, v. 49, n. 3, p. 199–203, 2007. Citado na página 28.
- GRIFFITH, T. Chaos and fractals in vascular biology. *Vascular Medicine Review*, SAGE Publications Sage UK: London, England, n. 2, p. 161–182, 1994. Citado na página 47.

- GROSS, J. L.; YELLEN, J.; ZHANG, P. *Handbook of graph theory*. [S.l.]: CRC press, 2013. Citado na página 29.
- HAUSDORFF, F. Dimension und äußeres maß. *Mathematische Annalen*, Springer, v. 79, n. 1-2, p. 157–179, 1918. Citado 2 vezes nas páginas 47 e 51.
- HOPKINS, B.; WILSON, R. J. The truth about königsberg. *The College Mathematics Journal*, Taylor & Francis, v. 35, n. 3, p. 198–207, 2004. Citado na página 28.
- HSU, L.-H.; LIN, C.-K. *Graph theory and interconnection networks*. [S.l.]: CRC press, 2008. Citado na página 32.
- IRWIN, M. D.; HUGHES, H. L. Centrality and the structure of urban interaction: measures, concepts, and applications. *Social Forces*, The University of North Carolina Press, v. 71, n. 1, p. 17–51, 1992. Citado na página 37.
- JEONG, H. et al. Lethality and centrality in protein networks. *Nature*, Nature Publishing Group, v. 411, n. 6833, p. 41–42, 2001. Citado na página 45.
- JOHN, C.; ALLAN, H. D. *A first look at graph theory*. [S.l.]: Allied Publishers, 1995. Citado na página 32.
- JONES, B. J. et al. Scaling laws in the distribution of galaxies. *Reviews of Modern Physics*, APS, v. 76, n. 4, p. 1211, 2005. Citado na página 47.
- KENETT, D. Y.; PERC, M.; BOCCALETTI, S. Networks of networks—an introduction. *Chaos, Solitons & Fractals*, Elsevier, v. 80, p. 1–6, 2015. Citado na página 47.
- KLEINBERG, J. *The small-world phenomenon: An algorithmic perspective*. [S.l.], 1999. Citado 2 vezes nas páginas 24 e 59.
- KLEINBERG, J. M. Navigation in a small world. *Nature*, Nature Publishing Group, v. 406, n. 6798, p. 845, 2000. Citado 3 vezes nas páginas 24, 59 e 62.
- LAM, C.-H.; SANDER, L. M. Fractals in surface growth with power-law noise. *Journal of Physics A: Mathematical and General*, IOP Publishing, v. 25, n. 3, p. L135, 1992. Citado na página 51.
- LIU, S. Fractals and their applications in condensed matter physics. In: *Solid State Physics*. [S.l.]: Elsevier, 1986. v. 39, p. 207–273. Citado na página 51.
- MANDELBROT, B. B. Fractals: form, change and dimension. *San Francisco: WH Freeman and Company*, 1977. Citado na página 47.
- MANDELBROT, B. B.; FRAME, M. Fractals. *Encyclopedia of physical science and technology*, Academic Press Orlando, FL, v. 5, p. 579–593, 1987. Citado na página 47.
- MARCUS, D. *Graph theory: a problem oriented approach*. [S.l.]: Maa, 2008. Citado na página 29.
- MASON, O.; VERWOERD, M. Graph theory and networks in biology. *IET systems biology*, IET, v. 1, n. 2, p. 89–119, 2007. Citado na página 36.

- MATHIEU, F.; COMTE, C. Kleinberg's grid unchained. *Theoretical Computer Science*, v. 826–827, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0304397518305966>>. Citado 2 vezes nas páginas 24 e 64.
- MATSUSHITA, M.; MEAKIN, P. Cluster-size distribution of self-affine fractals. *Physical Review A*, APS, v. 37, n. 9, p. 3645, 1988. Citado na página 51.
- MEAKIN, P. *Fractals, scaling and growth far from equilibrium*. [S.l.]: Cambridge university press, 1998. v. 5. Citado na página 47.
- MONTEIRO, L. H. A. Sistemas dinâmicos complexos. *Editores Livraria da Física, São Paulo*, 2010. Citado na página 31.
- NEWMAN, M. E.; BARABÁSI, A.-L. E.; WATTS, D. J. *The structure and dynamics of networks*. [S.l.]: Princeton university press, 2006. Citado na página 29.
- NEWMAN, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.*, American Physical Society, v. 89, p. 208701, Oct 2002. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.89.208701>>. Citado na página 35.
- NONNENMACHER, T. F.; LOSA, G. A.; WEIBEL, E. R. *Fractals in biology and medicine*. [S.l.]: Birkhäuser, 2013. Citado na página 47.
- PASTOR-SATORRAS, R.; VÁZQUEZ, A.; VESPIGNANI, A. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.*, American Physical Society, v. 87, p. 258701, Nov 2001. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.87.258701>>. Citado na página 35.
- PEEBLES, P. J. The fractal galaxy distribution. *Physica D: Nonlinear Phenomena*, Elsevier, v. 38, n. 1-3, p. 273–278, 1989. Citado na página 47.
- PHILLIPS, J. D.; SCHWANGHART, W.; HECKMANN, T. Graph theory in the geosciences. *Earth-Science Reviews*, Elsevier, v. 143, p. 147–160, 2015. Citado na página 33.
- PICKOVER, C. A. *Chaos and Fractals: A computer graphical journey*. [S.l.]: Elsevier, 1998. Citado na página 49.
- PIETRONERO, L. The fractal structure of the universe: Correlations of galaxies and clusters and the average mass density. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 144, n. 2-3, p. 257–284, 1987. Citado na página 47.
- RAHMAN, M. S. et al. *Basic graph theory*. [S.l.]: Springer, 2017. Citado na página 32.
- READ, R. C. *Graph theory and computing*. [S.l.]: Academic Press, 2014. Citado na página 33.
- REDNER, S. How popular is your paper? an empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 4, n. 2, p. 131–134, 1998. Citado na página 45.
- ROBERSON, M. R.; BEN-AVRAHAM, D. Kleinberg navigation in fractal small-world networks. *Phys. Rev. E*, v. 74, p. 017101, 2006. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.74.017101>>. Citado 4 vezes nas páginas 24, 63, 69 e 74.

- SCHANK, T.; WAGNER, D. Approximating clustering coefficient and transitivity. *Journal of Graph Algorithms and Applications*, Brown University, Providence, RI; University of Texas, Dallas, v. 9, n. 2, p. 265–275, 2005. Citado na página 33.
- SPIELMAN, D. A. Spectral graph theory and its applications. In: IEEE. *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*. [S.l.], 2007. p. 29–38. Citado na página 33.
- STEEN, M. V. Graph theory and complex networks. *An introduction*, v. 144, 2010. Citado na página 32.
- TRAVERS, J.; MILGRAM, S. The small world problem. *Sociometry*, v. 1, p. 425–443, 1967. Disponível em: <<https://doi.org/10.2307/2786545>>. Citado 4 vezes nas páginas 23, 55, 58 e 59.
- TRINAJSTIC, N. *Chemical graph theory*. [S.l.]: Routledge, 2018. Citado na página 30.
- TRUDEAU, R. J. *Introduction to graph theory*. [S.l.]: Courier Corporation, 2013. Citado na página 29.
- TURCOTTE, D. Fractals and fragmentation. *Journal of Geophysical Research: Solid Earth*, Wiley Online Library, v. 91, n. B2, p. 1921–1926, 1986. Citado na página 47.
- TURCOTTE, D. L. Fractals in geology and geophysics. *Pure and applied Geophysics*, Springer, v. 131, n. 1, p. 171–196, 1989. Citado na página 47.
- VASUDEV, C. *Graph theory with applications*. [S.l.]: New Age International, 2006. Citado na página 33.
- VICSEK, T. s. *Fractal growth phenomena*. [S.l.]: World scientific, 1992. Citado na página 51.
- VOSS, R. F. Random fractals: characterization and measurement. In: *Scaling phenomena in disordered systems*. [S.l.]: Springer, 1991. p. 1–11. Citado na página 50.
- WALLIS, W. D. *A beginner's guide to graph theory*. [S.l.]: Springer Science & Business Media, 2010. Citado na página 30.
- WATTS, D.; STROGATZ, S. Collective dynamics of 'small-world' networks. *Nature*, v. 393, p. 440–442, 1998. Disponível em: <<https://doi.org/10.1038/30918>>. Citado 6 vezes nas páginas 23, 33, 45, 56, 59 e 60.
- WATTS, D. J. *Six degrees: The science of a connected age*. [S.l.]: WW Norton & Company, 2004. Citado 2 vezes nas páginas 56 e 59.
- WEISSTEIN, E. W. Königsberg bridge problem. <https://mathworld.wolfram.com/>, Wolfram Research, Inc., 2002. Citado na página 28.
- WILSON, R. E.; GOSLING, S. D.; GRAHAM, L. T. A review of facebook research in the social sciences. *Perspectives on psychological science*, Sage Publications Sage CA: Los Angeles, CA, v. 7, n. 3, p. 203–220, 2012. Citado na página 44.
- YFANTIS, E.; FLATMAN, G.; ENGLUND, E. Simulation of geological surfaces using fractals. *Mathematical geology*, Springer, v. 20, n. 6, p. 667–672, 1988. Citado na página 49.

ZILL, D.; SHANAHAN, P. *A first course in complex analysis with applications*. [S.l.]: Jones & Bartlett Learning, 2009. Citado na página [48](#).

Apêndices



Relação entre as Distribuições Binomial e Poisson

No capítulo 2 discutimos as redes totalmente aleatórias e seus aspectos estatísticos. Abordamos o fato de que uma distribuição de Poisson pode ser obtida de uma distribuição binomial através de aproximação, com foco exclusivo nos benefícios operacionais aplicáveis a ciência de rede. Neste anexo fornecemos maiores detalhes matemáticos.

Seja uma probabilidade numa distribuição binomial dada pela [Equação A.1](#)

$$p(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (\text{A.1})$$

Para redes com muitos vértices ($N \rightarrow \infty$) a aproximação $\langle k \rangle = p(N-1) \simeq pN$ é válida, note que para manter $\langle k \rangle$ constante é necessário que $p \rightarrow 0$. Então vamos considerar os casos onde $k \ll N$, nessas condições o primeiro termo da [Equação A.1](#) pode ser aproximado por

$$\binom{N-1}{k} = \frac{(N-1)(N-2)\cdots(N-k)}{k!} \approx \frac{(N-1)^k}{k!}. \quad (\text{A.2})$$

Expandimos em série o logaritmo natural¹ do último termo

$$\ln[(1-p)^{(N-1-k)}] = (N-1-k)\ln(1-p) = (N-1-k)\ln\left(1 - \frac{\langle k \rangle}{N}\right)$$

¹ $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots, \forall |x| \leq 1.$

$$\ln[(1-p)^{(N-1-k)}] \approx (N-1-k) \frac{\langle k \rangle}{N-1} = -\langle k \rangle \left(1 - \frac{k}{N-1}\right) \approx -\langle k \rangle.$$

$$e^{\ln[(1-p)^{(N-1-k)}]} \approx e^{-\langle k \rangle}.$$

$$(1-p)^{(N-1-k)} \approx e^{-\langle k \rangle}. \quad (\text{A.3})$$

Substituindo as equações [A.2](#) e [A.3](#) na [A.1](#) temos

$$p(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} = \frac{(N-1)!}{k!} p^k e^{-\langle k \rangle} = \frac{(N-1)!}{k!} \frac{\langle k \rangle^k}{(N-1)!} e^{-\langle k \rangle}$$

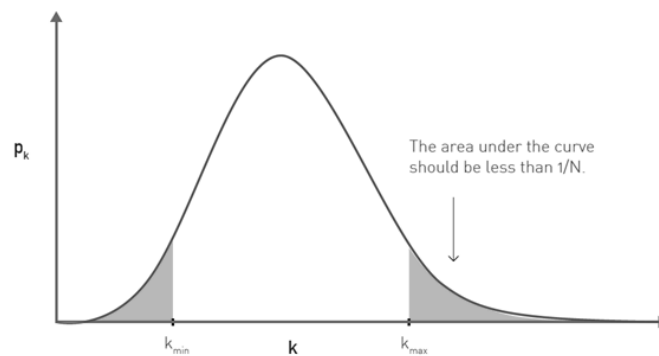
Que nos leva a distribuição de Poisson da [Equação A.4](#).

$$p(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}. \quad (\text{A.4})$$

Valores máximo e mínimo do grau médio $\langle k \rangle$ de um vértice

Em ciência de redes, define-se corte natural superior, denotado por k_{max} , o maior grau que um vértice pode manifestar em conformidade com a distribuição que rege a sua aleatoriedade. Estamos particularmente interessados na probabilidade de uma rede com N vértices apresentar apenas um nó com grau k_{max} .

Figura 37 – Distribuição de Poisson do grau k .



Fonte: (BARABÁSI, 2013)

Analisando graficamente a distribuição de Poisson, isso representa toda a área hachurada com $k \geq k_{max}$ na Figura 37. Uma vez que a distribuição probabilidade acumulada é $1 - P(k_{max})$, temos

$$1 - P(k_{max}) = 1 - e^{-\langle k \rangle} \sum_{k=0}^{k_{max}} \frac{\langle k \rangle^k}{k!} = e^{-\langle k \rangle} \sum_{k=k_{max}+1}^{\infty} \frac{\langle k \rangle^k}{k!},$$

$$1 - P(k_{max}) \approx e^{-\langle k \rangle} \frac{\langle k \rangle^{k_{max}+1}}{(k_{max} + 1)!}. \quad (\text{B.1})$$

Para uma rede de teste com $N = 10^9$ e $\langle k \rangle = 1000$, além do fato de que

$$N[1 - P(k_{max})] \approx 1.$$

Obtemos um corte natural superior de $k_{max} = 1185$.

Através de processo análogo, definimos e calculamos o corte natural inferior k_{min} levando em conta que

$$P(k_{min} - 1) = e^{-\langle k \rangle} \sum_{k=0}^{k_{min}-1} \frac{\langle k \rangle^k}{k!},$$

e

$$NP(k_{min} - 1) \approx 1.$$

Resolvendo para a rede de teste que adotamos encontramos $k_{min} = 816$.



Modelo de Kleinberg: Tempo de entrega para $\alpha = 2$

Em uma busca em redes de pequeno mundo utilizando o algoritmo descentralizado, dizemos que a mensagem está na fase j quando a distância entre o portador da mensagem e o alvo for $2^j \leq d \leq 2^{j+1}$. Note que a menor e a maior fase são respectivamente iguais a 0 e $\ln N$ e se queremos calcular o tempo médio de entrega T , é necessário determinar o tempo médio T_j em que a mensagem transita na fase j . Fazemos isso calculando a probabilidade P_{out} de a mensagem sair da fase j saltando para um nó $v \in \gamma$ através de um link de longo alcance. Sabendo que dentro de γ o número de nós N_γ cresce proporcionalmente ao quadrado do raio, temos que $N_\gamma \sim 2^{2j}$ e que cada um deles é uma fonte potencial de um link de longo alcance.

A distância entre u e qualquer nó de γ deve ser tal que $d(u, v \in \gamma) \leq 2^{j+1} + 2^j \leq 2^{j+2}$, então, usando a [Equação 4.1](#) determinamos P_{out}

$$P_{out} \sim \frac{N_\gamma}{(2^{j+2})^2 \ln N} \sim \frac{2^{2j}}{(2^{j+2})^2 \ln N} \sim \frac{1}{\ln N}. \quad (\text{C.1})$$

Note que se um dos nós de γ adquirir link de longo alcance com u a fase j é finalizada, em outras palavras, em cada passo dentro da fase j existe uma probabilidade proporcional a $1/\ln N$ de a mensagem saltar para outra fase mais próxima do alvo. Então, para a mensagem trafegar i passos na fase j serão necessários $i - 1$ fracassos na tentativa de sair de j e apenas um sucesso.

$$T_j(i) = \sum_{i=1}^{\infty} i Pr[T_j] \quad (\text{C.2})$$

$$T_j(i) = 1.Pr[T_j = 1] + 2.Pr[T_j = 2] + 3.Pr[T_j = 3] + \dots \quad (\text{C.3})$$

Aplicando algumas propriedades da estatística podemos escrever

$$T_j(i) = Pr[T_j \geq 1] + Pr[T_j \geq 2] + Pr[T_j \geq 3] + \dots \quad (\text{C.4})$$

Considerando os $(i - 1)$ insucessos, argumentamos que

$$Pr[T_j \geq i] \leq \left(1 - \frac{1}{\ln N}\right)^{i-1} \quad (\text{C.5})$$

substituindo a [Equação C.5](#) na [Equação C.4](#) temos

$$Pr[T_j \geq i] \leq 1 + \left(1 - \frac{1}{\ln N}\right) + \left(1 - \frac{1}{\ln N}\right)^2 + \left(1 - \frac{1}{\ln N}\right)^3 + \dots \quad (\text{C.6})$$

Que é uma PG com fator multiplicativo $\left(1 - \frac{1}{\ln N}\right)$ onde a soma converge para

$$\frac{1}{1 - \left(1 - \frac{1}{\ln N}\right)} = \ln N$$

Assim

$$T_j[i] \sim \ln N \quad (\text{C.7})$$

Dessa forma, reunindo todos os tempos de entrega computados nas $\ln N$ fases, temos:

$$T = \sum_{j=0}^{\ln N} T_j \sim \ln^2 N. \quad (\text{C.8})$$