

Grafos de Mundos Pequenos e Difusão de Inovações

Guilherme Steinberger Elias

Trabalho Final de Mestrado Profissional

Instituto de Computação
Universidade Estadual de Campinas

Grafos de Mundos Pequenos e Difusão de Inovações

Guilherme Steinberger Elias

Campinas, 16 de janeiro de 2005

Banca Examinadora :

- Prof. Dr. Jacques Wainer (Orientador)
Instituto de Computação – UNICAMP
- Prof. Dr. Jaime Simão Sichman
Escola Politécnica – USP
- Prof. Dr. Siome Klein Goldenstein
Instituto de Computação – UNICAMP
- Prof. Dr^a. Ariadne Maria B. Rizzoni Carvalho (Suplente)
Instituto de Computação – UNICAMP

UNIDADE	BC
Nº CHAMADA	T/UNICAMP EL42g
V	EX
TOMBO BC/	66709
PROC.	16-123-06
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	11,00
DATA	12/1/03
Nº CPD	

EL42g - 374499

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Elias, Guilherme Steinberger

El42g Grafos de mundos pequenos e difusão de inovações / Guilherme
Steinberger Elias -- Campinas, [S.P. :s.n.], 2005.

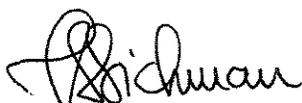
Orientador : Jacques Wainer

Trabalho final (mestrado profissional) - Universidade Estadual de
Campinas, Instituto de Computação.

1. Arquitetura de redes. 2. Redes neurais. 3. Algoritmos. 4. Teoria
dos grafos. I. Wainer, Jacques. II. Universidade Estadual de Campinas.
Instituto de Computação. III. Título.

TERMO DE APROVAÇÃO

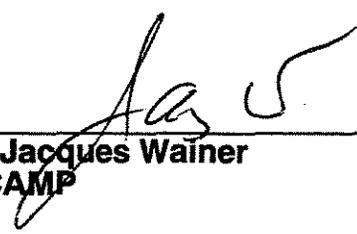
Trabalho Final Escrito defendido e aprovado em 24 de fevereiro de 2005, pela
Banca Examinadora composta pelos Professores Doutores:



Prof. Dr. Jaime Simão Sichman
ESCOLA POLITÉCNICA - USP



Prof. Dr. Siome Klein Goldenstein
IC - UNICAMP



Prof. Dr. Jacques Wainer
IC - UNICAMP

© Guilherme Steinberger Elias, 2005.
Todos os direitos reservados.

Agradecimentos

Agradeço ao Prof. Dr. Jacques Wainer pela orientação atenciosa e motivadora.

À minha esposa Anna Paula Ferreira por ter me dado força nos momentos mais trabalhosos e ter compreendido a importância desta atividade.

À minha mãe Margarethe Steinberger Elias por sempre ter estado ao meu lado e por me inspirar à pesquisa.

A Deus por ter me dado esta valiosa oportunidade.

Resumo

Esta é uma pesquisa, na área de Tecnologias de Informação, sobre como as inovações são disseminadas em sociedade. Tomando o trabalho de Everett Rogers (1962) sobre “difusão de inovações” como referência, o estudo situa-se na interface desta área com os campos da Sociologia e do Marketing, permitindo avaliar como as mudanças ou os novos conceitos são recebidos ou rejeitados em agrupamentos de diferentes naturezas. A representação de como as inovações penetram nestas populações tem sido feita através de simulações baseadas em grafos, onde cada vértice representa um membro da população. Nos estudos de Rogers e outros pesquisadores, prevalece uma forma aleatória de interconexão destes membros, aqui chamada de “mundo aleatório”.

Entretanto, mais recentemente, Watts (1999) descobriu que, sob certas condições, membros de grupos interconectam-se de forma que todos possuam apenas uns poucos intermediários entre si, em contextos onde haja vizinhanças com muitos membros já interconectados. Ainda hoje se estuda as causas para este fenômeno, chamado de “mundo pequeno”. O objetivo deste trabalho é verificar como se comporta a curva de “difusão de inovações” em contextos com tais características e identificar as diferenças em relação ao modelo onde as interconexões eram aleatórias.

Os primeiros resultados mostraram que, na maioria dos casos aqui analisados, a velocidade de aceitação de inovações (performance) é inferior em contextos de “mundo pequeno”, se comparada aos de “mundo aleatório”. Ao redesenharmos, entretanto, modelos de “mundo pequeno” com as mesmas definições de Watts e extraindo a aleatoriedade, encontramos uma performance superior, ainda que bem mais sensível do que as versões de Rogers e Watts em relação às características da população.

Concluímos que a adoção de alguns modelos de “mundo pequeno” (desde que bem parametrizados) pode ser mais produtiva e permitir preservar características do mundo real, ou seja, obter efeito mais “natural”. Testamos diferentes parâmetros, encontrando como mais produtivos: o grau de independência das novas conexões em relação às anteriores (parâmetro α), a probabilidade aleatória de conexão entre dois membros quaisquer (parâmetro p), a quantidade média de conexões por membro (parâmetro km), a seleção de membros inovadores como líderes locais (além de possuírem conexão com outros grupos, detêm alto índice de conexões com seu próprio grupo). O modelo para simulação aqui utilizado, com pequeno ajuste, foi o de “multi-agentes” idealizado por Maienhofer e Finholt (2000).

Abstract

This is a research in the area of Information Technology about how are the innovations disseminated in society. Taking the work of Everett Rogers (1962) about “diffusion of innovations” as reference, the study takes place in the interface of this area with the fields of Sociology and Marketing, allowing evaluating how do the changes or new concepts are received or rejected in groups of different natures. The representation of how do the innovations penetrate in these populations has been done by simulations based on graphs, where each vertex represents one member of the population. In the studies of Rogers and other researchers, prevails a random shape of interconnection of these members, here known as “random world”.

But recently, Watts (1999) discovered that, under certain conditions, members of groups interconnect in such a way where all of them have only a few intermediates between each other, in contexts where there are neighborhoods with many members already connected. Still today, it's studied the causes for this phenomenon, known as “small world”. The objective of this work is to verify how is the behavior of the curve of “diffusion of innovations” in contexts with these characteristics and identify the differences regarding the model where the interconnections were random.

The first results showed that, in the majority of the cases here analyzed, the velocity of acceptance of innovations (performance) is lower in contexts of “small world”, if compared to those from “random world”. But, after redesigning models of “small world” based on the same definitions of Watts and extracting the randomness, we found a superior performance, even though it's more sensible than versions of Rogers and Watts regarding the characteristics of the population.

We concluded that the adoption of some models of “small world” (if with good parameters), can be more productive and allow the preservation of characteristics of real world, that means, to obtain a more “natural” effect. We tested different parameters, finding advantages in the following: the degree of independence of new connections regarding the previous connections (parameter α), the random probability of connection between any two members (parameter p), the medium quantity of connections per member (parameter km), the selection of innovators members as local leaders (they have connection with other groups and high index of connections with their own group). The model used here for simulation, with little adjust, was the “multi-agents” idealized by Maienhofer and Finholt (2000).

Sumário

Resumo.....	8
Abstract	9
Lista de Figuras	11
Lista de Tabelas.....	12
Capítulo 1	13
Introdução	13
1.1 Fundamentos de “difusão de inovações”	16
1.2 Fundamentos de “mundo pequeno”	31
1.3 Trabalhos relacionados com o tema.....	34
Capítulo 2	37
Representação dos modelos utilizando grafos.....	37
2.1 Grafos d-Lattice	40
2.2 Grafos aleatórios	42
2.3 Grafos do tipo “connected caveman”.....	43
2.4 Grafos “mundo pequeno”	45
2.5 Resumo de L e χ	56
Capítulo 3	57
Modelo para estudo prático	57
3.1 Características dos modelos de grafos utilizados	57
3.2 Funcionalidades do simulador	61
Capítulo 4	75
Resultados do estudo prático	75
4.1 Análise de Regressão Logística.....	75
4.2 Resultados e análise	78
4.3 Análise de Teste-t.....	91
Capítulo 5	93
Conclusões	93
5.1 Desdobramentos futuros desta pesquisa.....	95
Apêndice A.....	96
Cálculo de L em um grafo “Connected Caveman”	96
Bibliografia.....	100

Lista de Figuras

Figura 1 - Exemplo de curva em forma de “S”.....	22
Figura 2 - Processo de decisão para uma inovação.....	24
Figura 3 - Distribuição dos adotantes com relação ao seu grau de inovação	26
Figura 4 – Processo de desenvolvimento da inovação.....	27
Figura 5 – Variáveis que determinam a taxa de adoção de inovações.....	29
Figura 6 – Exemplo de grafo.....	37
Figura 7 – Exemplo de grafo 1-lattice com $km=2$	40
Figura 8 – Exemplo de grafo 1-lattice com $km=4$	41
Figura 9 – Exemplo de grafo aleatório.....	42
Figura 10 – Exemplo de grafo do tipo “isolated caveman”.....	43
Figura 11 – Exemplo de grafo “connected caveman”.....	44
Figura 12 – Exemplo de grafo aleatório.....	45
Figura 13 – Interface do software de simulação.....	62
Figura 14 – Grafo aleatório obtido no simulador.....	63
Figura 15 – Adoção ao longo do tempo - grafo aleatório.....	64
Figura 16 – Grafo aleatório com substrato obtido no simulador.....	65
Figura 17 – Adoção ao longo do tempo - grafo aleatório com substrato....	66
Figura 18 - Grafo “mundo pequeno” com substrato – modelo α	67
Figura 19 - Adoção ao longo do tempo - grafo “mundo pequeno” com substrato – modelo α	68
Figura 20 - Grafo “mundo pequeno” adaptado com substrato.....	69
Figura 21 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado com substrato.....	70
Figura 22 -Grafo “mundo pequeno” adaptado sem substrato.....	71
Figura 23 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado sem substrato.....	72
Figura 24 - Grafo “mundo pequeno” adaptado sem substrato com 8 grupos	73
Figura 25 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado sem substrato com 8 grupos.....	74
Figura 26 – Tipos de vértices do Grafo aproximado do “connected caveman”.....	96

Lista de Tabelas

Tabela 1 – Comparativo da clusterização e largura de caminho característico para Distribuição de Energia no oeste dos EUA	54
Tabela 2 – Comparativo da clusterização e largura de caminho característico para sistema nervoso da formiga <i>C. Elegans</i>	55
Tabela 3 – Resumo do cálculo da clusterização e largura de caminho característico para alguns tipos de grafos	56
Tabela 4 – Resultados com parâmetros de referência	80
Tabela 5 – Resultados com aumento da quantidade de nós	81
Tabela 6 – Resultados com aumento da quantidade de nós e de k	82
Tabela 7 – Resultados com aumento de n e de inovadores	83
Tabela 8 – Resultados com aumento do <i>threshold</i> de adoção mínimo	84
Tabela 9 – Resultados com menor <i>threshold</i> de adoção máximo	85
Tabela 10 – Resultados com maior intervalo entre <i>thresholds</i> de adoção mínimo e máximo	86
Tabela 11 – Resultados com maior influência dos vizinhos adotantes.....	87
Tabela 12 – Resultados com aumento da memória de cada nó	88
Tabela 13 – Resultados comparando-se com os modelos de “mundo pequeno” criados manualmente.....	90
Tabela 14 – Resultados do Teste-t	91

Capítulo 1

Introdução

Neste trabalho, relacionado com os campos de Computação, Sociologia e Marketing, pesquisaremos através de simulações computacionais, como as inovações são disseminadas em diferentes modelos de sociedade (um modelo tradicional e um recentemente descoberto). Toda idéia ou mudança, pode ser considerada como uma **inovação** por indivíduos com potencial de aceitação (aqui chamados de “potenciais adotantes”), de acordo com a forma como ela os afetará. A percepção do indivíduo perante esta novidade irá definir se ela é ou não uma inovação, e o seu estado de adoção ou rejeição pode ser neutro, mesmo que ela já tenha sido apresentada há bastante tempo. Portanto, não é apenas o grau de novidade que a inovação representa para o indivíduo que irá definir seu impacto no sistema. Existem outras variáveis, como, por exemplo, o grau de persuasão dos vizinhos potenciais adotantes.

Exemplos de inovações :

- Ferver ou filtrar a água para evitar doenças causadas pelos germes;
- Layout das teclas no teclado (DVORAK x QWERTY);
- Telefone celular;
- Sistema métrico;
- Fertilizantes e agrotóxicos;
- Cinto de segurança em automóveis;
- Proibição de se fumar durante vôos.

Uma inovação surge da percepção de um problema ou necessidade. Esta fase é objeto de pesquisas sobre como suprir esta necessidade. Tais pesquisas desenvolvem-se através de uma análise das tecnologias ou recursos existentes capazes de permitir o seu desenvolvimento. Se a solução idealizada for viável, um agente gerador de mudanças irá selecioná-la e iniciar o seu desenvolvimento ou planejamento. Em seguida, ela será divulgada por um inovador e irá desencadear um processo de difusão desta inovação.

A forma como a inovação é introduzida em uma população e suas consequências, é o objeto de estudo de uma subárea das Tecnologias da Informação conhecida como **difusão de inovações**. As mudanças ou os novos conceitos podem ser aceitos ou rejeitados em agrupamentos de diferentes naturezas.

Utilizaremos o trabalho de Everett Rogers (1962) sobre “difusão de inovações” como referência. Para a representação de como as inovações penetram nas populações, utilizaremos simulações baseadas em grafos, onde cada vértice representa um membro da população. O modelo para simulação aqui utilizado, com pequeno ajuste, foi o de “multi-agentes” idealizado por Maienhofer e Finholt (2000).

Nos modelos utilizados por de Rogers e outros pesquisadores, prevalece uma forma aleatória de interconexão entre os membros, aqui chamada de “mundo aleatório”.

Recentemente, Watts (1999) descobriu que esta interconexão entre os membros não é tão aleatória como representada nos modelos de Rogers e outros pesquisadores. Sob certas condições, membros de grupos de diversas naturezas, encontrados no mundo real, interconectam-se de forma que todos possuam apenas uns poucos intermediários entre si, em contextos onde haja vizinhanças com muitos membros já interconectados. Este fenômeno é chamado de “mundo pequeno” ou “small world” em inglês. O objetivo deste trabalho é o de verificar como se comportará a “difusão de inovações” em modelos com tais características e identificar as diferenças em relação ao modelo onde as interconexões eram aleatórias.

Os estudos de “difusão de inovações” começaram na Europa no início do século XX, com o nascimento da Sociologia como ciência social. Por volta de 1962, Everett M. Rogers lançou um livro intitulado “Diffusion of Innovations”, onde reuniu e formalizou os principais conceitos de Difusão de inovações até aquele momento, os quais estaremos explicando no próximo capítulo. Desde então, intensificou-se o aparecimento de artigos e bibliografias nesta área e este autor tornou-se a principal referência de “difusão de inovações” até hoje.

Por volta da década de 60, outros estudos iniciaram-se motivados pela curiosidade em entender como certas populações conseguem possuir comportamentos sincronizados e harmônicos que as beneficiam sem que todos os indivíduos se comuniquem entre si.

Fenômenos tais como o de centenas de vaga-lumes piscando em sincronia perfeita também serviram de inspiração para esse tipo de investigação. Sabendo-se que cada vaga-lume possui isoladamente pouca inteligência e que os vaga-lumes machos

procuram piscar sincronizadamente para atrair as fêmeas, é interessante investigar como eles conseguem se sincronizar tão perfeitamente.

Entre outros fenômenos inspiradores, existe também o de centenas de grilos produzindo sons perfeitamente sincronizados durante o verão ou o de milhares de células do coração enviando sinais elétricos para o resto do coração de forma a sincronizar perfeitamente as suas contrações. Ou ainda, o de milhões de neurônios gerando pulsos elétricos para gerar funções cerebrais de percepção, ou mesmo o de centenas de pessoas batendo palmas sincronizadamente.

Como se formam estes comportamentos coletivos extremamente sincronizados e frequentemente vantajosos para as populações sem que exista um guia ou maestro é uma das questões mais intrigantes, tanto biológica quanto matematicamente. São fenômenos que indicam uma tendência da natureza em se organizar.

A recente conclusão do mapeamento dos genes humanos pelo projeto Genoma surpreendeu ao comprovar que o ser-humano possui pouco mais de trinta mil genes, que é pouco mais do que a quantidade de genes de algumas plantas, indicando que a complexidade não está intimamente relacionada com a quantidade de genes, pode estar relacionada com a forma como estes genes estão configurados (cf. Buchanan, 2002).

Outro fator motivador, segundo Watts (2003), é a pequena quantidade de intermediários entre qualquer par de indivíduos da população humana do planeta, em torno de seis para uma população de 6 bilhões. O estudo destes fenômenos abrange diversas disciplinas desde Biologia e Matemática até Sociologia, conhecido como “mundo pequeno”. Para estes estudos, o autor de referência é Duncan J. Watts (cf. 1999 e 2003), que os introduziu publicou obra em co-autoria com Steven Strogatz (cf. Watts e Strogatz, 1994), um estudioso de fenômenos não lineares e de sistemas dinâmicos. Strogatz foi o mentor de Watts no início de suas pesquisas.

Neste trabalho, verificaremos de forma prática, a performance da “difusão de inovações” (Rogers, 1983) nos modelos de “mundo aleatórios” (utilizados nas simulações de “difusão de inovações” de Rogers e outros pesquisadores) e nos de “mundo pequeno” (Watts, 1999). Em seguida identificaremos os pontos mais relevantes. Esperamos que os resultados possam, de alguma forma, auxiliar nos estudos para aplicações de grafos de “mundo pequeno” em diversas áreas e no aperfeiçoamento dos estudos de “difusão de inovações”. Um exemplo de aplicação seria no auxílio ao profissional de marketing para introduzir mais rapidamente (time to speed) um produto no mercado.

No capítulo 1, apresentaremos os conceitos básicos de “difusão de inovações” (ROGERS, 1983 e 2003) e de “mundos pequenos” (WATTS, 1999 e 2003), assim como uma breve revisão da literatura relacionada ao tema. Apresentaremos também alguns trabalhos relacionados com este tema. No capítulo 2, expomos alguns grafos utilizados nos estudos destes fenômenos, o que permite iniciar, no capítulo 3, um estudo prático com simulações, que permitirá identificar as características relevantes nos modelos de agrupamento em análise. Tal tarefa mostrará os comportamentos de populações perante a inserção de inovações através de diferentes abordagens de arquiteturas de conexões. No capítulo 4, analisaremos os resultados destas simulações através de um comparativo entre os modelos utilizados. No capítulo 5, sistematizaremos as principais conclusões que esta pesquisa deixou entrever.

1.1 Fundamentos de “difusão de inovações”

Na época em que se iniciaram os estudos sobre “difusão de inovações”, por volta do início do século XX, o francês Gabriel Tarde, um dos pais da Sociologia e Psicologia Social, tratou de forma analítica as tendências de sua sociedade, encontrando certas generalizações na “difusão de inovações” que publicou em seu livro “Leis da imitação” (apud Rogers, 1983). Entre as motivações de Tarde para a pesquisa nesse campo, estava a de entender por que 90% de diferentes inovações de um mesmo período eram esquecidas, enquanto outros 10% obtinham sucesso.

Outro estudioso tomado como referência na subárea de estudos de “difusão de inovações” foi o alemão Georg Simmel, que havia estudado Filosofia e foi um dos primeiros sociólogos de sua época, por volta de 1908. Simmel confrontou-se com o conceito do “estranho” - um indivíduo membro do sistema, mas pouco vinculado a ele, o que lhe permitia tornar-se um inovador e desviar-se mais facilmente das normas do sistema social no qual estava inserido.

Outros estudos ocorreram na mesma época, principalmente na Inglaterra, Alemanha e Áustria e, posteriormente, também nos Estados Unidos. Pesquisas formais sobre “difusão de inovações” iniciaram-se de forma independente nas décadas de 40 e 50, sobretudo nas áreas agrícola e acadêmica. Elas chamaram a atenção por obterem resultados equivalentes, por exemplo, com relação à curva de adotantes de inovações ao longo do tempo, cujo formato obtido era em “S”. Além disso, comprovavam que, em uma população de adotantes, os inovadores possuíam status sócio-econômico superior

ao dos últimos adotantes de uma inovação. Um exemplo do modelo de curva em “S” segue-se na Figura 1, explicada mais adiante.

Em 1962, Everett M. Rogers tornou-se famoso por sua descrição e formalização em detalhes de um modelo de difusão capaz de dar mais consistência às pesquisas tradicionais da época (Rogers, 1983). Desde então, os estudos sobre o assunto têm se intensificado. Rogers assinalou que as áreas que mais publicaram sobre a “difusão de inovações” foram, em ordem decrescente, a de Sociologia Rural (20%), Marketing/Administração (16%) e Comunicação (15%).

Uma das motivações para o estudo de “difusão de inovações” é o fato de que a adoção de uma inovação por uma população é freqüentemente difícil e demorada, mesmo quando é óbvio que uma tal inovação geraria grandes vantagens. Portanto, um dos desafios na área é o de estudar os critérios que fazem aumentar a velocidade desta adoção.

1.1.1 Conceitos do estudo de Difusões de Inovações em um sistema social

Os conceitos aqui expostos foram obtidos de Rogers (1983 e 2003), por ser considerado a principal referência em “difusão de inovações”.

No estudo da difusão de inovações, define-se o **sistema social** como um conjunto de unidades Inter-relacionadas, engajadas em atividades diversas, para obter um objetivo comum que as mantém unidas. Seus membros podem ser indivíduos, grupos informais, organizações ou outros subsistemas. Entre os componentes estudados dentro do sistema social, destacam-se as estruturas sociais, cuja existência justifica-se porque as unidades não possuem comportamentos idênticos entre si, existindo padrões regulares entre parcelas de unidades. Houve poucos estudos sobre a influência na adoção ou rejeição de inovações pela estrutura social, em comparação com outros tópicos da pesquisa da difusão. Um exemplo de pesquisa foi a realizada em 1981 por Rogers e Kincaid com duas mulheres coreanas que traziam as seguintes características em comum: analfabetas, casadas, cada uma com dois filhos, ambas com 29 anos de idade, com maridos de mesma escolaridade e com fazendas do mesmo tamanho. Ao contrário do esperado, ambas tiveram comportamentos de adoção diferentes com relação a uma mesma inovação de métodos anticoncepcionais. Isso ocorreu, provavelmente, porque elas viviam em ambientes sociais diferentes. A vila A onde

morava uma delas adotou em 57% a inovação enquanto que a vila B onde morava a outra adotou em 26% a mesma inovação. Estas vilas possuíam estruturas sociais diferentes com relação ao uso de métodos anticoncepcionais, pois na vila B o planejamento familiar era religiosamente mal visto. Os moradores da vila A devem ter tido mais influência na primeira mulher a favor do uso de anticoncepcionais do que os da vila B para a segunda mulher.

Outro componente utilizado nos estudos de um sistema social é o das **normas** que o organizam, que podem ser definidas como padrões de comportamentos tolerados estabelecidos para os membros de um mesmo sistema social. Indicam para os indivíduos qual o comportamento que se espera deles, e podem ser barreiras para a entrada de inovações benéficas no sistema.

Um componente que pode ser associado com o desenvolvimento de muitas inovações é a **tecnologia**. Em geral, ela possui uma parte física, como, por exemplo, semicondutores, e uma base de informações indicando como ela deve funcionar ou como deve ser utilizada. Um exemplo de informações contidas nesta base é o uso de instruções codificadas que definem o seu comportamento. A dosagem entre estes componentes pode variar muito.

Existem situações em que algumas inovações estão relacionadas entre si, onde a adoção de uma induz a adoção da outra, o que facilita a sua aceitação. Este tipo de abordagem é mais realista do que a abordagem da análise de inovações individuais e é conhecida como **tecnologia em cluster**¹. Um exemplo ocorreu durante a Segunda Guerra Mundial, quando houve uma “Revolução Verde” nos países do Terceiro Mundo com o aparecimento simultâneo de fertilizantes químicos e agrotóxicos. Aqueles que adotassem ambas as tecnologias conseguiam até mesmo triplicar sua produção de arroz e vendê-la para os países em guerra.

A facilidade de se ajustar uma inovação para que seja aplicada para outros fins diferentes dos planejados é um fator importante a ser considerado na sua durabilidade e sucesso de adoção. Isto é conhecido como **reinvenção**. Um exemplo é a flexibilidade de ajustes do sistema operacional Linux, que permitem, por exemplo, a criação de soluções otimizadas de softwares para armazenamento, roteadores de rede, eletroeletrônicos como *palmtops* e até mesmo injeções eletrônicas de automóveis.

¹ Cluster em grafos representa agrupamentos que se destacam na população.

Quando dois ou mais membros de um sistema social criam e compartilham informações entre si com o objetivo de atingir um entendimento mútuo, temos o processo de **comunicação**. Este processo pode ocorrer de forma bidirecional ou unidirecional, gerando ou não uma convergência com relação a como tratar certos eventos. Este processo de interação costuma ocorrer ao longo de vários ciclos. Exemplo: um agente inovador tenta persuadir seus clientes a adotarem uma inovação. Outro exemplo : um cliente apresenta uma necessidade ou problema ao agente inovador e este indica uma possível solução.

Outra característica analisada no estudo de “difusão de inovações” é a forma como as mensagens chegam entre os indivíduos, conhecida como **canal de comunicação**. Ele influi na facilidade de transmissão das informações da inovação para os potenciais adotantes, assim como no seu grau de persuasão. Alguns exemplos são os canais de comunicação em massa, geralmente os mais rápidos, eficientes e de maior credibilidade, como, por exemplo, a televisão, jornais, rádio e Internet. No entanto os meios de comunicação de massa possuem menos persuasão do que canais interpessoais com indivíduos de um mesmo nível sócio-econômico e educacional com condições de estabelecer um contato face a face². As investigações na área de difusão têm demonstrado que a maioria dos que adotam a inovação (**adotantes**) não valida cientificamente as inovações que absorvem e, sim, observam como outros indivíduos conhecidos estão se beneficiando desta adoção. Portanto, isso sugere que o processo central de difusão é um modelamento e imitação desenvolvido pelos potenciais adotantes com relação aos seus conhecidos que já adotaram a inovação.

Quando a novidade atinge um indivíduo do sistema social, ele terá um certo grau de dificuldade para escolher entre possíveis alternativas com relação ao seu tratamento, pois o potencial adotante fica em dúvida se a inovação será realmente útil. Este grau de dificuldade é tratado como **incerteza**.

A idéia de que inovações geram vantagens em relação ao que já está em uso, impele o indivíduo a entender melhor sobre a novidade e, assim, as dúvidas sobre a inovação serão minimizadas. Ele poderá, então, decidir entre adotá-la ou não. Por conseguinte, a dificuldade no tratamento de uma inovação pode ser diminuída através de mais informações. Entre estas informações, existem atributos que influenciam na percepção da inovação, como por exemplo à **vantagem relativa**. Ela pode ser vista

² O grau de similaridade entre dois ou mais indivíduos interagindo com relação a certos atributos como interesses, crença, educação e status sócio-econômico é conhecido como “homofilia”.

como sendo o grau de vantagem medido através de fatores como economia, conveniência e satisfação. O importante é o quanto os indivíduos percebem a inovação como sendo vantajosa, pois isto influenciará diretamente na velocidade da sua adoção. Os fatores econômicos se destacam na vantagem relativa. Se o preço inicial da inovação for atraente, a sua taxa de adoção tenderá a ser maior e será melhor ainda, se houver avanços que permitam diminuições no seu custo de produção ao longo do tempo. Outro fator que exerce influência é o status social obtido com a adoção da inovação.

Uma informação importante para o potencial adotante, durante a sua escolha, é o grau de consistência da inovação com relação aos valores anteriores e normas do seu sistema social, conhecida como **compatibilidade**. Isto influenciará diretamente na velocidade da sua adoção. Um exemplo de falta de compatibilidade é a adoção de métodos anticoncepcionais em regiões aonde a religião vai contra o planejamento ou restrição da gravidez.

A **complexidade** ou grau de dificuldade de entendimento e uso da inovação, também influi diretamente na sua velocidade de adoção. Um exemplo é a dificuldade de entender a teoria dos germes como sendo a razão para se ferver ou filtrar a água.

Se o potencial adotante puder testar a inovação sem compromisso, suas chances de adoção deverão aumentar pois isso aumenta a confiança e ajusta as expectativas. Isto é conhecido como **testabilidade** e também influi na velocidade da sua adoção.

O grau de visibilidade ou **exposição** dos resultados da inovação é outro fator de influência na escolha e também está relacionado com a velocidade de adoção. Um exemplo é a adoção do uso de aparelhos de celulares, onde o potencial adotante vê com facilidade as vantagens que os adotantes conhecidos em seu meio estão tendo, pois é uma inovação utilizada em público.

A forma como é feita a decisão, entre adotar ou não a inovação, é outro fator de destaque que influencia no sucesso e velocidade de adoção. Existem decisões que são assumidas de forma **coletiva**, através de um consenso entre os membros do sistema, onde todas as unidades devem se conformar com a decisão depois que ela for tomada. As decisões tomadas pelo próprio indivíduo de forma independente com relação à opinião de outros membros do sistema são conhecidas como **opcionais**, embora a decisão ainda possa ter influência das normas do sistema ou de relacionamentos interpessoais. Este tipo de decisão costuma ser mais rápido do que as coletivas. Existem também as decisões **autoritárias**, onde a escolha é feita por poucos indivíduos do sistema que possuam poder, status ou grande conhecimento técnico. Os outros membros

simplesmente implementarão a decisão com pouca ou nenhuma influência na sua escolha. Em geral, geram uma taxa de adoção mais rápida do que as anteriores.

Um exemplo da forma como é assumida a decisão foi o uso do cinto de segurança nos automóveis. Inicialmente, o próprio proprietário do automóvel tinha que pagar a sua instalação (decisão opcional). Em seguida, uma lei federal determinada por um consenso entre os membros do congresso passou a vigorar, requerendo que o cinto de segurança já viesse instalado nos automóveis (decisão coletiva). Mais adiante, algumas leis estaduais chegaram a determinar que o uso do cinto de segurança seria obrigatório sob pena de multa (decisão autoritária).

Quando uma inovação é comunicada através de canais entre os membros de um sistema social ao longo do tempo, ocorre a **difusão**. Assim, ela pode ser compreendida como a forma como as alterações ocorrem dentro de uma estrutura de sistema social quando novos conceitos culturais ou tecnologias são inseridas. Pode ser considerada como um tipo de comunicação, onde o conteúdo está relacionado com inovações, gerando sempre um certo grau de incerteza quanto a sua aceitação. Exemplo: como indicar para uma população que se deve filtrar a água antes de bebê-la ? Como acontece uma revolução política ?

Os estudos sobre difusão de inovação tentam desvendar mistérios como as diferenças entre os adotantes iniciais e finais, o impacto de cada característica da inovação em sua aceitação e na sua velocidade de convergência, além da inesperada proeminente subida no percentual de adoção após atingir entre 10 e 25 % da população. As difusões que ocorrem de forma planejada e gerenciada são categorizadas sob o termo **disseminação**. Elas costumam ocorrer de forma centralizada, ou seja, a decisão sobre quando começar a difusão de uma inovação, quem deve aprová-la e por quem propagá-la é tomada por um pequeno número de agentes inovadores. Já as difusões espontâneas costumam ocorrer de forma descentralizada. Elas são compartilhadas pelos clientes potenciais e, portanto, existe uma estrutura mais horizontal sobre como será propagada a inovação. Em difusões extremamente descentralizadas pode até mesmo não existir o agente inovador, sendo os clientes potenciais com maior tendência de adoção os principais responsáveis pela difusão. Exemplo: novas idéias vindas da experiência de certos indivíduos pertencentes ao conjunto de clientes de uma empresa.

A **taxa de adoção** durante o processo de “difusão de inovações” representa a velocidade relativa com que uma inovação é assumida por membros de um sistema social. Geralmente, ela é medida através do número de indivíduos adotantes em um

certo período. Quando se coloca em um gráfico o número de indivíduos adotantes em uma frequência cumulativa ao longo do tempo, em geral obtém-se como resultado uma curva em forma de “S”, indicando que, no início, somente poucos indivíduos adotam a inovação (inovadores). Veja um exemplo na Figura 1 :

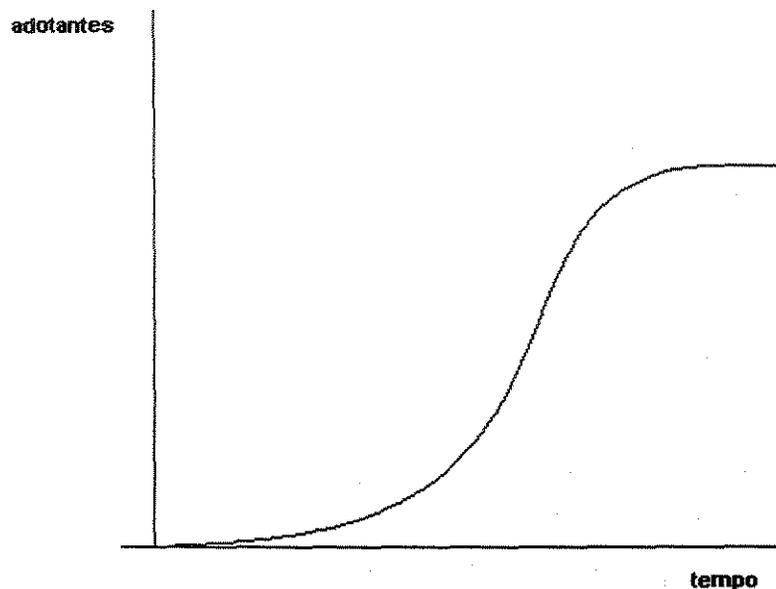


Figura 1 - Exemplo de curva em forma de “S”

Com o aumento de indivíduos adotantes, a curva começa a subir cada vez mais rápido e em seguida ela tende a se manter em um patamar, já quando faltam poucos potenciais adotantes. O *slope* (largura da curva) será menor para inovações rapidamente adotadas e maior para inovações com maior demora para adoção. Existem também diferenças no *slope* da taxa de adoção de uma mesma inovação causadas por diferenças entre os sistemas sociais onde ela é inserida. Estas diferenças podem ser, por exemplo, as leis, normas, fatores econômicos e educacionais. O número cumulativo de adotantes pode ser estimado porque a curva em “S”, em geral, é simétrica em relação ao seu ponto de inflexão. Se, depois deste ponto, a curva continuar em crescimento, haverá grande chance de sucesso na difusão.

O momento a partir do qual a difusão se torna auto-sustentável, sem a necessidade da atuação dos inovadores, é conhecido como ponto de massa crítica. Ela ocorre quando existe uma quantidade suficiente de indivíduos adotantes, de forma que a inovação continuará a ser adotada por outros indivíduos não adotantes, sem que sejam necessários esforços para isso. Ocorre a partir do ponto de inflexão da curva em “S” da adoção ao longo do tempo. Ele também pode representar o início da rejeição e descontinuidade da difusão.

O intervalo do período de decisão pelo uso de uma inovação, é geralmente medido a partir do momento do primeiro conhecimento obtido em seu processo, até o momento de decisão entre adotá-la ou rejeitá-la. Muitos divulgadores da inovação tentam aumentar a sua velocidade de adoção, através de uma comunicação mais eficiente. Mas o rápido conhecimento da existência da inovação pelos potenciais adotantes não garante que eles irão adotá-la, embora permita que se obtenha os resultados finais mais rapidamente.

Comparando a análise de Rogers da taxa de adoção ao longo do tempo com outras análises de estrategistas de marketing, percebemos que a principal diferença é que na área de marketing estima-se uma queda de adoção da tecnologia depois de atingido o patamar quase constante de adoção (Porter, 1980). Rogers comenta que o conceito de “reinvenção” pode fazer uma tecnologia permanecer mais tempo no mercado e não colocou foco na sua fase de declínio. A curva estimada pelos estrategistas de marketing também possui a forma em “S” e está dividida nas seguintes fases :

- introdução;
- crescimento;
- maturidade;
- declínio.

Um dos autores da área de marketing (Richards, 1997), mostrou uma curva interessante, indicando como um produto induziu a venda do outro ao longo do tempo. Este exemplo está diretamente relacionado com o conceito de tecnologia em *cluster* de Rogers, onde em um conjunto de inovações relacionadas entre si, a adoção de uma induz a adoção da outra.

O futuro da inovação após a sua difusão, será consequência da sua adoção ou rejeição ao longo do tempo. Dependerá, por exemplo, dos seus efeitos no sistema social terem sido funcionais ou disfuncionais; do fato de as mudanças geradas terem sido um resultado direto ou indireto da sua adoção; e também pelo seu reconhecimento como benéficas ou negativas e intencionais ou espontâneas.

Por conseguinte, podemos resumir o processo de decisão para uma inovação como sendo a inicial obtenção do conhecimento sobre a inovação por um indivíduo, unidade de tomada de decisão, organização ou comunidade, seguida por sua atitude com relação a adotá-la ou rejeitá-la. No caso de adoção, será feita a sua implementação e uso a fim de confirmar a decisão de continuar adotando-a, conforme a Figura 2 :

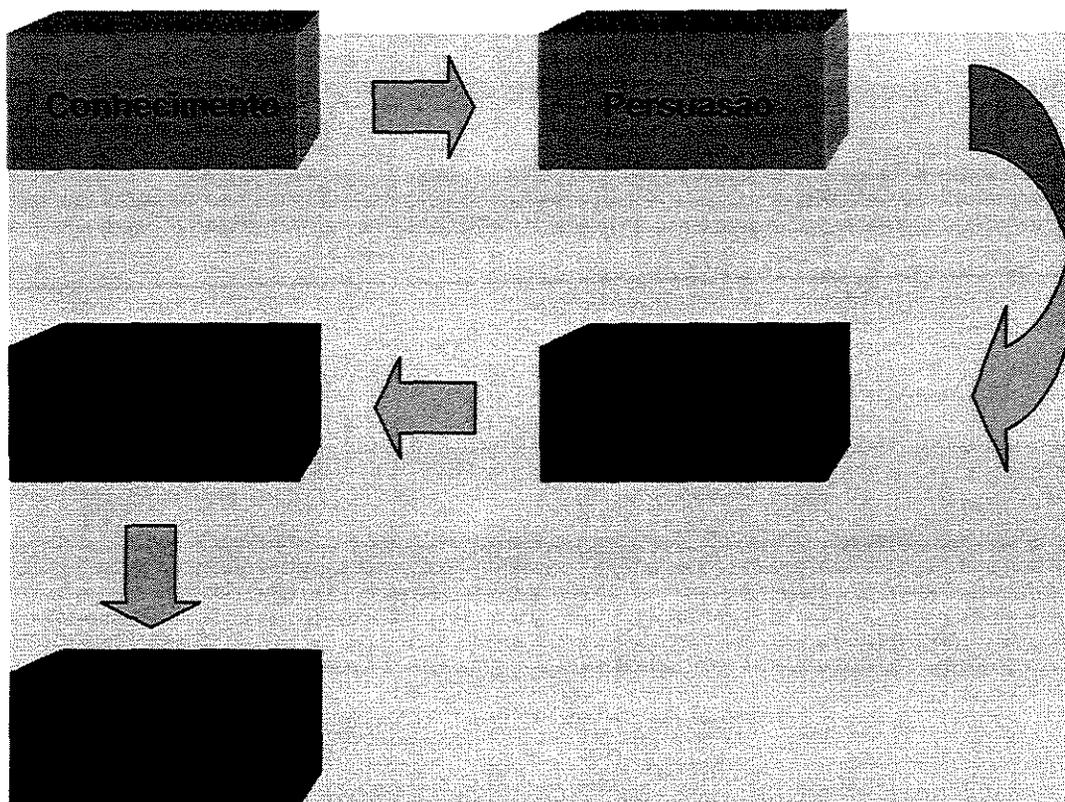


Figura 2 - Processo de decisão para uma inovação

1. Conhecimento : o indivíduo aprende sobre a existência da inovação e como ela funciona.
2. Persuasão : o indivíduo forma uma opinião favorável ou desfavorável com relação à inovação.

3. Decisão : o indivíduo executa atividades que o farão escolher entre adotar ou não adotar.
4. Implementação : o indivíduo faz o uso prático da inovação.
5. Confirmação : o indivíduo reforça ou não a sua decisão inicial de adoção.

1.1.2 Classificação do indivíduo com relação ao seu grau de inovação

Os indivíduos de um sistema social podem ser classificados com relação a sua possibilidade de adoção. Os indivíduos **inovadores**, por exemplo, são pessoas que sempre procuram informações sobre novas idéias, possuem maior exposição aos meios de comunicação em massa, sua rede de contatos interpessoais atinge sistemas distantes. Lidam bem com o grau de incerteza sobre as inovações por saberem lidar com complexidades técnicas. São sempre os primeiros a aderir, freqüentemente os responsáveis por validar as inovações. Também podem ser conhecidos como agentes de mudança.

Outro tipo de adotante é o **adotante precoce**, também conhecido como **formador de opinião**, que possui liderança na formação da opinião dos potenciais adotantes. Esta liderança não depende de status ou posição formal do indivíduo no sistema e, sim, da sua competência técnica, acessibilidade social e conformidade com as normas do sistema. Estão mais integrados ao sistema local do que os inovadores. São respeitados pelos potenciais adotantes e responsáveis pelo aumento da velocidade de adoção. Possuem nível educacional maior do que os que adotam a inovação mais tarde. Seu status sócio-econômico também costuma ser superior, e costumam utilizar as inovações como uma das formas de melhorá-lo.

Existe uma outra parcela da população que é classificada como a **maioria que adota mais cedo**. São indivíduos que provêm interconexões interpessoais que ajudam a difundir a inovação. Interagem mais com os seus vizinhos e conhecidos, representam cerca de 1/3 dos membros do sistema. Seu processo de tomada de decisão para adoção ou rejeição é mais demorado do que os dos mencionados anteriormente.

Outra parcela é representada pela **maioria que adota mais tarde**, composta por indivíduos mais céticos e cautelosos, que adotam geralmente por motivos econômicos ou por pressão de vizinhos que já adotaram e representam 1/3 dos membros do sistema.

Existe também a parcela dos **últimos adotantes**, que são indivíduos que possuem pouca liderança, poucos relacionamentos interpessoais em seu sistema. Só interagem com outros indivíduos céticos, precisam ter certeza de que a inovação não falhará e querem evitar gastos desnecessários.

Rogers verificou que a curva de distribuição dos adotantes ao longo das classificações mencionadas acima, em geral, seguem uma distribuição normal ao gerarem a curva de adoção em forma de “S”. Ele utilizou fatores como a média e o desvio padrão para dividir os adotantes ao longo desta distribuição normal, conforme o exemplo da Figura 3 :

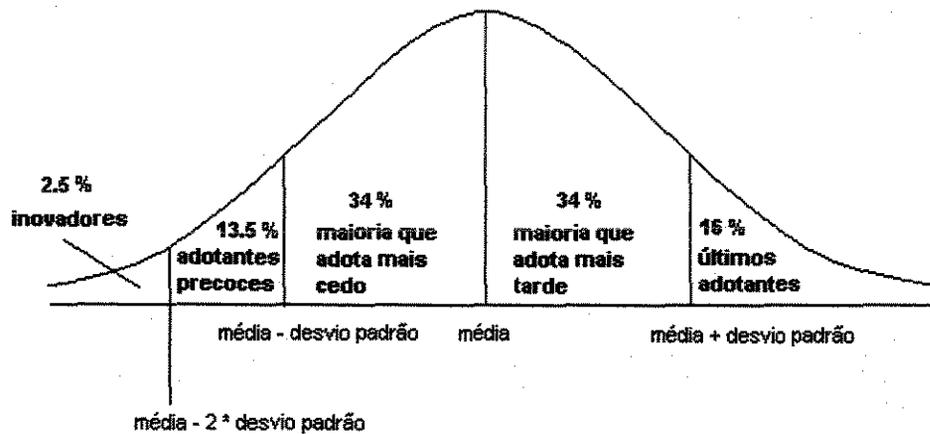


Figura 3 - Distribuição dos adotantes com relação ao seu grau de inovação

É importante mencionar que as percentagens mencionadas na Figura 3 são sugeridas por Rogers (1984 e 2003), mas que podem sofrer variações de acordo com as características de cada população que for estudada.

1.1.3 O processo de desenvolvimento da inovação

Existem diversas etapas no ciclo de vida de uma inovação. As decisões e atividades ocorridas antes do início do processo de difusão, exercem uma forte influência no futuro da sua evolução. Em geral, é necessário atuar científica e politicamente em todo o processo para que se desenvolva uma inovação com sucesso em sua difusão. As principais etapas podem ser vistas na Figura 4 :

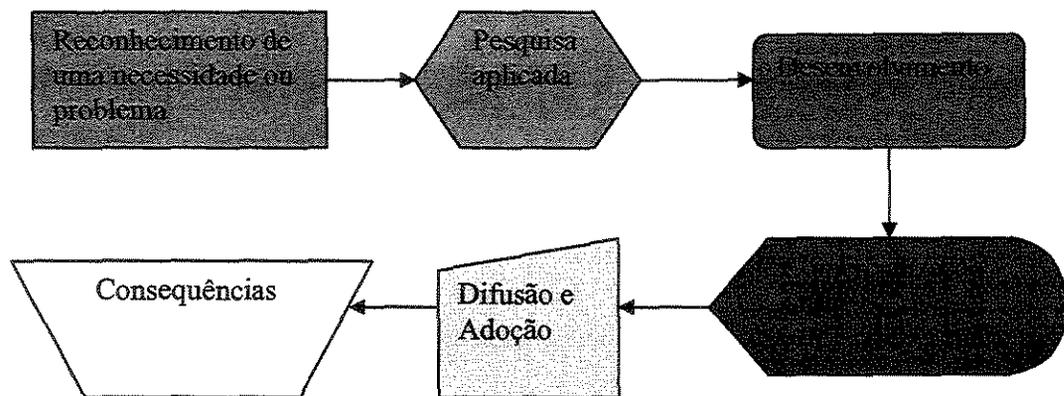


Figura 4 – Processo de desenvolvimento da inovação

1.1.4 Descrição das etapas do processo de desenvolvimento da inovação

- **Reconhecimento de uma necessidade ou problema** : reconhecer uma necessidade ou problema, estimulará a pesquisa e desenvolvimento de atividades necessárias para a criação de uma inovação que possa supri-la.
- **Pesquisa Aplicada** : Frequentemente, a inovação necessita de pesquisa básica para ser criada, considerada como um estudo científico e investigações sem fins específicos, mas que possam ser aproveitados. A pesquisa aplicada é o direcionamento deste estudo para a resolução de uma necessidade ou problema.
- **Desenvolvimento** : o desenvolvimento de uma inovação é o processo de se dar forma a uma idéia abstrata, esperando-se atender a uma necessidade de potenciais adotantes. Nesta etapa, o inventor deve procurar prever problemas que os últimos adotantes poderão encontrar com relação à inovação. O comportamento e regras de sua equipe, competidores, leis do governo, patentes existentes, entre outros, irão afetar o futuro da inovação e devem ser previstos.

- **Comercialização** : nesta etapa, a inovação pode ser considerada como um resultado científico, modelado de forma a estar pronta para ser adotada por usuários. Este processo costuma ser feito em empresas privadas. A comercialização é a produção, modelagem, marketing e distribuição de um produto que contém a inovação. Ela é a conversão de uma idéia de pesquisa em um produto ou serviço à venda no mercado.
- **Difusão e Adoção** : esta etapa é crucial no futuro da inovação. Neste momento, costuma-se ter pressões para que a inovação tenha a sua difusão aprovada rapidamente a fim de suprir logo as necessidades e problemas. Ao mesmo tempo, a reputação do divulgador estará dependendo das conseqüências benéficas que ocorram para os adotantes e, portanto, é um momento de cautela.
- **Conseqüências** : esta é a etapa final do processo de desenvolvimento da inovação, quando as mudanças geradas por ela para indivíduos ou sistemas sociais que a adotaram, serão aprovadas ou rejeitadas.

Segue na Figura 5 um diagrama representando algumas das principais variáveis que determinam a taxa de adoção de inovações :

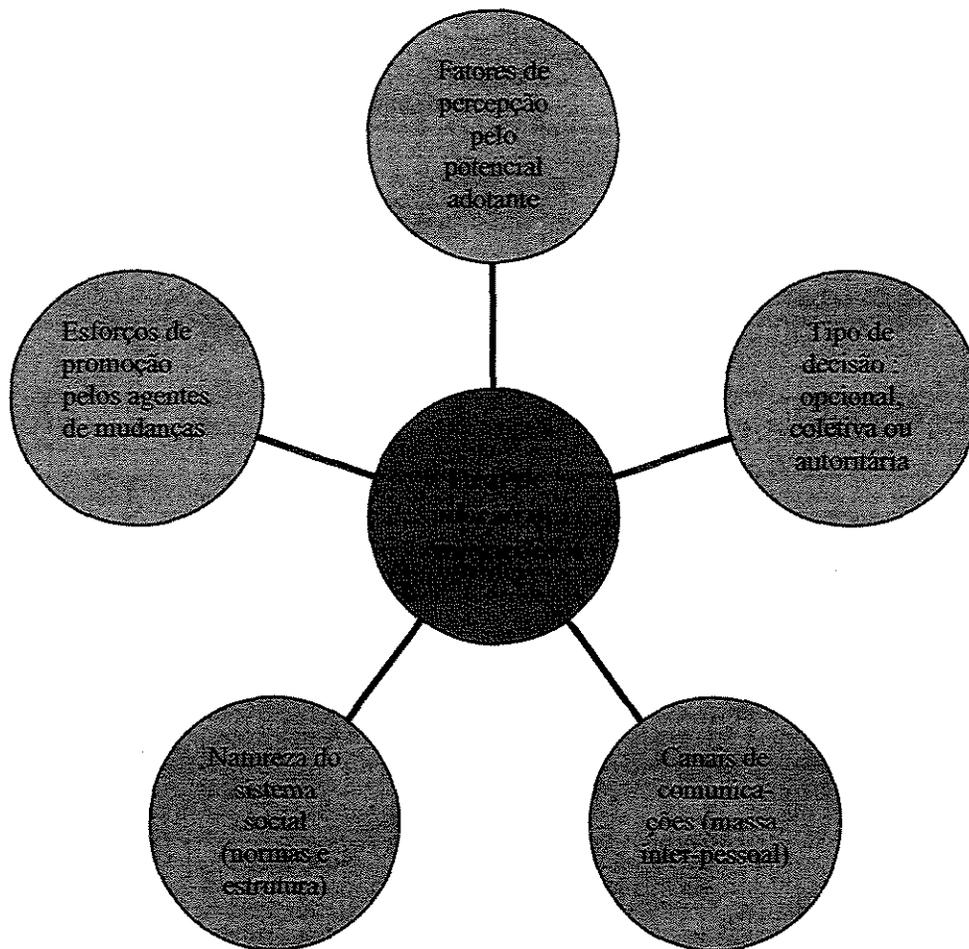


Figura 5 – Variáveis que determinam a taxa de adoção de inovações

1.1.5 Paradoxo entre necessidade da inovação e resistência em adotá-la

Geralmente os indivíduos ou organizações que mais necessitam dos benefícios de uma idéia costumam estar entre os últimos a adotá-la e os que a adotam mais cedo costumam ser os que menos necessitam dos seus benefícios. Um exemplo prático é a adoção do uso de métodos anticoncepcionais, onde os casais com melhor poder aquisitivo os utilizam freqüentemente embora tenham poder financeiro para ter vários filhos e costumem ter poucos filhos. Os casais com menor poder aquisitivo costumam ter mais filhos mesmo tendo menor poder financeiro e evitam o uso dos métodos anticoncepcionais.

Uma das razões deste paradoxo é o fato de que os agentes de mudanças costumam seguir uma estratégia por segmento da população ao difundirem as inovações, sempre escolhendo segmentos com menor resistência, que costumam ser as elites sócio-econômicas, as quais possuem maior conhecimento e capacidade de compreensão dos benefícios a serem obtidos. Este tipo de estratégia tende a aumentar a desigualdade social pois os benefícios demoram a chegar nas camadas com menor poder aquisitivo.

A estrutura social de uma população possui grandes influências sobre como ocorrerá a difusão. Estudaremos mais adiante como ocorre a difusão em diferentes estruturas sociais.

1.2 Fundamentos de “mundo pequeno”

O “mundo pequeno” é um tipo de estrutura social, onde para qualquer par de indivíduos origem-destino, o menor percurso para se conseguir uma comunicação terá muito poucos intermediários em relação a quantidade de indivíduos existentes na população. A quantidade de ligações na vizinhança de cada indivíduo é muito alta em relação ao seu máximo valor possível, ou seja, é muito comum que os vizinhos de um indivíduo se conheçam entre si.

Os conceitos aqui expostos foram obtidos de Watts (1999 e 2003), por ser considerado a principal referência em “mundos pequenos”.

Os estudos de “mundo pequeno” iniciaram-se em torno de 1960 com análises matemáticas elaboradas por Manfred Kochen e Ithiel de Sola Pool. Suas principais contribuições foram no cálculo do número médio de conhecidos por cada membro da população, e no cálculo da probabilidade de dois membros selecionados ao acaso estarem conectados com no máximo dois intermediários. Na mesma época, existiram publicações de Anatol Rapoport e Solomonoff, da Universidade de Chicago, sobre estatísticas de como as doenças se propagam em populações de acordo com as suas estruturas de conexões.

Em 1983, Granovetter (apud Watts, 1999) defendeu que os laços mais fracos entre membros da população são mais importantes do que os laços fortes na rede social pois em geral os laços mais fracos formam links importantes entre grupos distintos de conhecidos. Esta idéia foi reforçada por Skvoretz e Fararo em 1989, demonstrando que quanto maior a frequência de laços fracos, menos intermediários um membro escolhido ao acaso terá até qualquer outro membro da população.

Os últimos estudos sobre redes sociais têm sido na área de dimensionamento e geometria de espaço onde elas existem. Estes estudos concordam que cada membro de uma população possui coordenadas, que representam um conjunto de características suficientes para se distinguir um membro do outro de forma única. O observador conhece as distâncias entre cada par de membros, onde a distância é definida de acordo com o contexto analisado, mas ele não conhece as coordenadas pois não sabe a dimensão do espaço. A relação entre coordenadas e distâncias entre cada par de membros não é padronizada. Alguns criaram suas próprias metodologias de métricas, outros como Barnett, Pool e Kochen utilizaram metodologias sem métricas, alegando

que são necessárias para representar as redes sociais com precisão. Eles defendem, por exemplo, que se o membro “a” está muito próximo de “b” e “c”, então em sistemas com métricas, “b” e “c” também deveriam estar bem próximos entre si, mas na prática pode acontecer de “b” e “c” estarem muito distantes entre si. Estas interconexões só poderiam ser representadas por modelos sem métricas. De nosso ponto de vista, em geral não se deve utilizar coordenadas físicas e distâncias não físicas em uma mesma representação gráfica. No entanto, isto seria válido em casos específicos, onde se queira comparar os dois tipos de distâncias em um mesmo gráfico. Uma sugestão seria o uso de coordenadas físicas por indivíduo e distâncias medidas por quantidade de intermediários, onde membros com poucos intermediários entre si poderiam ter tons de cores próximos. O contexto das distâncias varia de acordo com o que se deseja representar. Um exemplo disto é o fato de que “a” e “b” podem estar muito próximos fisicamente e muito distantes em quantidade de intermediários entre si.

Uma das inspirações para o estudo deste fenômeno foi o jogo de Kevin Bacon, criado por um cientista da computação chamado Brett Tjaden, da Universidade de Virgínia. A idéia deste jogo era a de mapear a relação de conhecidos entre a maior parte dos atores e atrizes que já existiram no meio cinematográfico. O interessante neste jogo foi a percepção de que, embora Kevin Bacon fosse um ator que não costumava ser o personagem central nos 55 filmes em que participou, ele estava, de certa forma, no centro do universo cinematográfico no sentido de que a maioria dos atores que existiram no cinema o conheceram direta ou indiretamente. Depois, descobriu-se que isto não ocorria somente com este ator. Isto foi mensurado do seguinte modo :

- Se os atores ou atrizes tivessem participado do mesmo filme que Kevin, teriam o número de Bacon como sendo “1”;
- Se nunca estiveram em um filme com o Kevin, mas estiveram em um filme com algum outro ator que já esteve no mesmo filme de Kevin, então eles teriam o número de Bacon como sendo “2”, e assim por diante.

O cientista Brett, defendeu que o número de Bacon não passou de “4” para quem participou de filmes americanos. Elvis Presley, por exemplo, obteve um número de Bacon como sendo “2”. Esta teoria foi confirmada através de uma base de filmes

existente na Internet no site IMDb³ e ajudou a motivar o tema sobre o fenômeno “mundo pequeno”.

Outras motivações para o estudo do “mundo pequeno” são as freqüentes “coincidências” - pessoas estranhas entre si conversam em diversos eventos que ocorrem pelo mundo e, na maioria dos casos, encontram inesperadamente algum conhecido em comum e exclamam : “Sim, o mundo é pequeno !”

A primeira experiência empírica sobre o assunto foi feita por Stanley Milgram em 1967 (apud WATTS,1999). Ele enviou uma série de cartas registradas, cuja entrega podia ser monitorada, para intermediários em Kansas e Nebraska. As cartas tinham instruções pedindo que fossem reenviadas para um único destinatário em Massachusetts, mas só descreviam algumas características deste destinatário como localização aproximada, profissão e demografia. Os intermediários só poderiam reenviar as cartas para conhecidos que tivessem maior probabilidade de conhecer o destinatário. O resultado da experiência indicou uma corrente de até 6 intermediários para alcançar o destinatário, causando grande surpresa pois construir conscientemente correntes tão pequenas que existissem para qualquer membro da população já seria um trabalho cujo cálculo é extremamente difícil.

O nosso alcance social no mundo é limitado ao nosso grupo de conhecidos diretos, sendo que vários destes membros conhecem-se entre si. Além disso, este grupo representa uma parcela extremamente pequena da população global (no máximo milhares comparados com bilhões), o que torna o fenômeno do “mundo pequeno” ainda mais interessante.

Desde o experimento de Milgram, a maior parte dos trabalhos que surgiram têm se empenhado em determinar características em grupos sociais como :

- Número de intermediários entre membros;
- Número de conhecidos de cada membro;
- A estrutura dos grupos relacionando o círculo de conhecidos de um membro com os de outros membros.

As principais dificuldades encontradas nestes trabalhos têm sido :

- Obter informações sobre quem são os conhecidos de cada membro é difícil em grandes populações;

³ O site de internet do IMDb fica no endereço <http://www.us.imdb.com>.

- Na prática, os laços de cada membro com seus conhecidos não possuem a mesma importância e isto é difícil de ser representado;
- A conectividade entre os membros não é sempre simétrica, ou seja, os subordinados costumam se considerar conectados aos seus superiores mas o contrário nem sempre ocorre;
- O conceito de “conhecido” é interpretado de formas diferentes de acordo com a idéia ou informação a ser transmitida. Exemplo: Os conhecidos para quem se pode divulgar “boatos” nem sempre são os mesmos para quem se pode pedir ajuda.

As pessoas geralmente estimam incorretamente a quantidade de conhecidos que possuem e isso gera dificuldades para se modelar corretamente às conexões entre os indivíduos.

Neste trabalho serão utilizados grafos para representarem mundos pequenos e mundos aleatórios. Em seguida serão estudadas as vantagens de um sob o outro com relação a performance de difusões de inovações.

Estaremos também estudando em que condições os elementos de uma rede conectada conseguem estar próximos entre si.

1.3 Trabalhos relacionados com o tema

Durante a elaboração deste trabalho, foram pesquisados outros estudos relacionados com o mesmo assunto. Não foi encontrado nenhum trabalho propondo o mesmo objetivo aqui exposto, no entanto, foram encontrados artigos interessantes, um deles com resultados que reforçam os que foram aqui obtidos.

O artigo “Finding Optimal Targets for Change Agents: A computer Simulation of Innovation Diffusion”, de Maienhofer e Finholt (2000), analisa a performance de “difusão de inovações” em “mundos aleatórios”. Isto é feito através de um modelo para simulação da difusão ao longo do tempo, cujas características parecem relevantes e úteis para a simulação em outros tipos de populações. Por isso este modelo foi adotado na elaboração do simulador deste trabalho. Neste artigo os autores sugerem o uso de

variáveis como *threshold*⁴ de rejeição ou de adoção, e o controle de memória por nó que permite simular a influência das tendências de adoção sob os indivíduos. A curva de adoção, gerada pelo modelo proposto neste artigo, obteve a esperada curva em “S” mencionada por Rogers (1983 e 2003). Os autores também inseriram em seu modelo a capacidade de desadoção da inovação, que não se aproveitou neste trabalho para não perder o foco da análise comparativa entre “difusão de inovações” no “mundo aleatório” e no “mundo pequeno”. Nas análises do modelo de seu trabalho, os autores perceberam que a seleção criteriosa dos nós como inovadores influencia no sucesso de adoção da inovação pela população em “mundos aleatórios”. Além disso perceberam que pequenas alterações nos thresholds e parâmetros de influência dos vizinhos ou variações de adoção causaram grandes alterações no resultado final. Isto também foi percebido nos resultados do trabalho aqui exposto, porém em graus diferentes de acordo com o tipo de população utilizada.

Outro artigo interessante é o “The Small-world phenomenon: an algorithmic perspective”, de Kleinberg (1999). Neste artigo, o autor afirma que, nenhum algoritmo descentralizado onde um indivíduo qualquer possua somente informações da sua vizinhança conseguirá gerar “*short paths*”⁵ com outros membros da população sem probabilidade de fracasso. Kleinberg executa simulações em uma grande quantidade de variações dos modelos criados por Watts (1999), e consegue encontrar em um deles uma alta probabilidade de que indivíduos com informações locais consigam criar os “short paths”. Ele consegue isso através de determinados níveis de correlação entre a estrutura local de conexões de cada indivíduo e as conexões de longa distância.

Goldberg e Roth (2002) apresentam mais uma evidência do fenômeno “mundo pequeno” na natureza, utilizando o modelo de “mundos pequenos” para auxiliar na predição de possíveis conexões entre proteínas (não documentadas experimentalmente). Os autores conseguiram acertar quatro conexões conhecidas em treze tentativas, sendo que só se conhecem cerca de 1200 de 10^7 possíveis, o que indica que o fenômeno também parece existir nas conexões entre as proteínas.

⁴ Threshold pode ser interpretado como um valor que limita dois estados de uma característica ou atributo.

⁵ Short paths significa uma quantidade acima de dois intermediários entre dois membros de uma população na ausência da aresta entre eles.

Mihajlo, Fred, Annexstein e Berman (2001) explicam alguns conceitos sobre “*power law*”⁶ em grafos “mundo pequeno”. O artigo menciona que distribuições do tipo “*power law*” indicam a função da variação de uma característica de um nó em relação ao ranking desta na população elevado a uma potência representada por uma constante : $y = x^k$. Esta propriedade pode ser, por exemplo, o nível de *threshold* de um nó, a sua frequência de variação, número de nós na vizinhança para uma distância específica, entre outras.

Amaral, Scala, Barthélémy e Stanley (2000) utilizam o conceito de distribuição com regra de potência (“*power law*”), para explicar a tendência de novas conexões em diferentes modelos de “mundo pequeno”. Lago-Fernández, Huerta, Corbacho e Sigüenza1 (1999) comparam a propagação da atividade em neurônios utilizando “mundo aleatório” e “mundo pequeno”. Obtêm como resultado uma melhor performance do “mundo aleatório”, que é o mesmo resultado obtido inicialmente no trabalho aqui exposto. No entanto, a maior rapidez de difusão no “mundo aleatório” compromete a coerência do comportamento dos neurônios e os autores concluem que o “mundo pequeno” também possui uma performance ainda considerada boa e mantendo esta coerência. O texto não apresenta detalhes de como foram gerados os seus modelos de “mundo aleatório” e “mundo pequeno”.

⁶ Power law são regras de potência em matemática, onde uma variável depende da variação de uma outra elevada a uma potência.

Capítulo 2

Representação dos modelos utilizando grafos

Para conseguirmos fazer simulações de difusões de inovações em populações conectadas através de diferentes arquiteturas, utilizamos modelos baseados em grafos. Os grafos são apropriados para simulações em que seja relevante a forma como os elementos estão conectados entre si. Veja um exemplo na Figura 6.

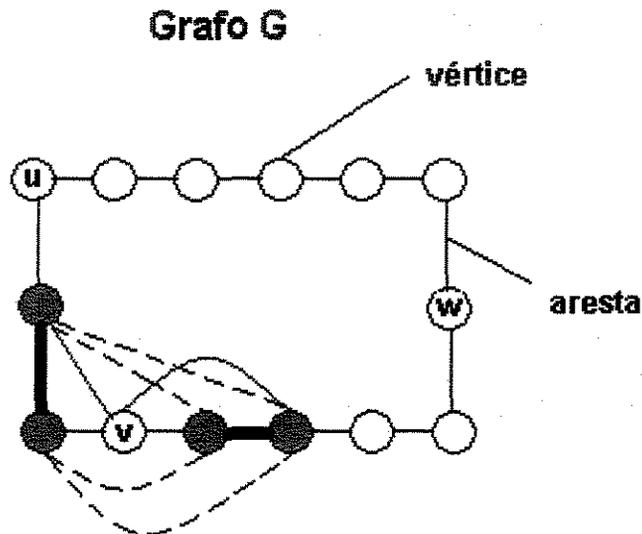


Figura 6 – Exemplo de grafo

Componentes importantes aqui utilizados :

- **Grafo G** : conjunto de elementos chamados vértices e de conexões entre estes elementos, chamadas de arestas;
- **Conjunto de vértices $V(G)$** : conjunto total de vértices que existem no grafo G;

- **Ordem n do grafo G** : quantidade de vértices $V(G)$. No exemplo da Figura 6 temos $n=14$;
- **Tamanho M do grafo G** : quantidade de arestas $E(G)$. Na Figura 6 temos $M=16$;
- **Conjunto de arestas $E(G)$** : conjunto total de arestas que existem no grafo G;
- **Grau médio k_m do grafo** : representa a quantidade média de conexões por vértice em todo o grafo. Na Figura 6 temos $k_m=32/14=2.286$;
- **Quantidade de conexões k** : representa a quantidade de conexões de um vértice que pertença ao grafo. Na Figura 6 temos $k=4$ para o vértice v, representado pelas linhas vermelhas;
- **Vizinhança do vértice $\Gamma(v)$** : são os vértices adjacentes ao vértice v. No exemplo da Figura 6, os vértices com a cor azul representam a vizinhança do vértice v;
- **Clustering γ_v** : Mede o quanto os vértices adjacentes ao vértice v são adjacentes entre si :

$$\gamma_v = \frac{|E(\Gamma(v))|}{\binom{k}{2}}$$

ou

$$\gamma_v = \frac{|E(\Gamma(v))|}{(k * \frac{(k-1)}{2})}$$

onde $|E(\Gamma(v))|$ é o número de arestas na vizinhança do vértice v (sem considerar suas próprias arestas) e no denominador temos a combinação dois a dois de cada vizinho do vértice v (número de possíveis arestas na vizinhança do vértice v, sem considerar suas próprias arestas). No exemplo da Figura 6 teríamos que o clustering somente do vértice v seria de duas arestas existentes na sua vizinhança

(arestas em negrito) de um total de seis arestas possíveis (arestas em negrito + arestas pontilhadas), ou seja, 0,333;

- **Número de Bacon j** : representa a quantidade mínima de vértices intermediários entre 2 vértices quaisquer. No exemplo da Figura 6 temos que $j=3$ entre os vértices v e w ;
- **Largura de caminho característica $L(G)$** : representa a média da quantidade mínima de arestas entre quaisquer 2 vértices.
- **Diâmetro D do grafo** : máximo número de Bacon no grafo. Um grafo com topologia em anel com n vértices e $k = 2$ tem por exemplo $D = \frac{n}{2}$.

Na Figura 6 temos $D=5$ (atingido por exemplo entre os vértices u e w);

Uma das formas de se garantir que seja gerado um grafo conectado, sem inserir mais estrutura do que o necessário e sem perder as propriedades a serem estudadas, é a da inicialização da construção do grafo com um **substrato** conectado, sob o qual as novas arestas serão adicionadas. Para evitar uma distorção nas propriedades estudadas, este substrato deverá possuir a menor estrutura homogênea conectada necessária, evitando assim que algum elemento tenha maior influência do que os outros.

A única estrutura que satisfaz esta condição é o “anel”, onde cada vértice possui exatamente duas arestas (Watts,1999). Além disso, ele permite uma largura de caminho característica L maximizada em relação a outros substratos conectados como, por exemplo, estrelas e árvores.

Veremos a seguir alguns exemplos de grafos extremamente organizados e desorganizados, utilizados como extremos de referência para a análise de outros grafos com relação à largura de caminho característica e clusterização. Estes grafos foram utilizados por Watts durante suas tentativas em gerar “grafos de mundo pequeno”, portanto eles serão mencionados utilizando Watts (1999) como referência.

2.1 Grafos d-Lattice

São grafos extremamente organizados, de dimensão d , onde qualquer vértice v é conectado aos seus k vizinhos laterais, com $k_m \geq 2 * d$, onde k_m é a quantidade média de conexões de cada vértice.

Veja exemplos de grafos d-lattice nas figuras abaixo :

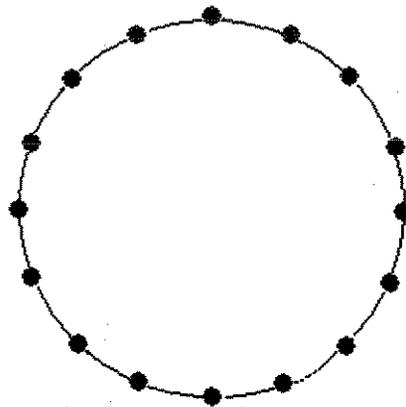


Figura 7 – Exemplo de grafo 1-lattice com $k_m=2$

Neste exemplo, temos um grafo de uma dimensão onde cada nó possui duas conexões, formando um anel.

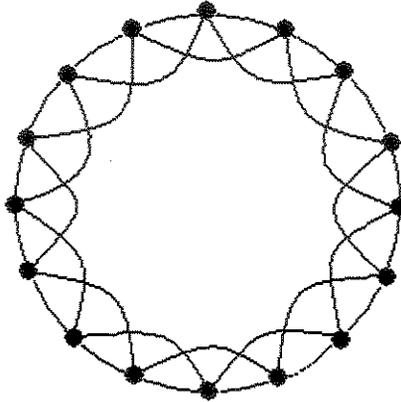


Figura 8 – Exemplo de grafo 1-lattice com $km=4$

Neste exemplo, temos um grafo de uma dimensão, onde cada nó possui quatro conexões.

A vantagem do uso de grafos d-lattice, é o fato de facilitar o cálculo da Largura de caminho característica e do Clustering. Assim podemos utilizá-los facilmente como referência para a análise de outros grafos.

Para um grafo do tipo 1-lattice com $km=2$, teríamos uma topologia de conexões em anel. Para um grafo 2-lattice com $km=4$, teríamos uma topologia de conexões em forma quadriculada.

Em grafos 1-lattice com $km \geq 2$ temos :

$$L = \frac{n * (n + km - 2)}{(2 * km * (n - 1))} \quad e \quad \gamma = 3 * \frac{(km - 2)}{4 * (km - 1)}$$

A quantidade de vértices com j ou menos intermediários até qualquer outro vértice será: $\Lambda_j = j * km$

2.2 Grafos aleatórios

É um exemplo de grafo extremamente desorganizado. Seu conjunto de arestas é gerado aleatoriamente. Watts chama este modelo de grafo de “Solaria”, onde não existe propensão de conexão entre elementos específicos.

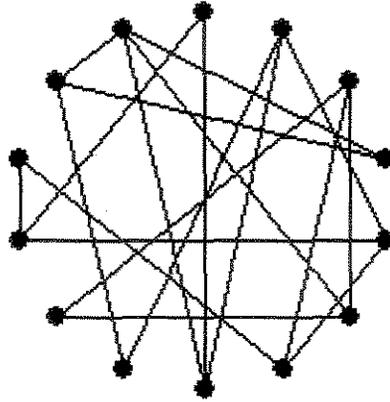


Figura 9 – Exemplo de grafo aleatório

Quase toda a teoria de grafos aleatórios baseia-se em dois modelos :

$G(n,M)$: possui um conjunto de vértices $V(G)$ com arestas M aleatoriamente escolhidas. É o modelo utilizado neste trabalho.

$G(n,p)$: possui um conjunto de vértices $V(G)$, onde cada uma das possíveis arestas de $\binom{n}{2}$ existe com probabilidade $0 < p < 1$ independente para cada aresta.

A partir do momento em que o grafo aleatório possui mais do que $\frac{n}{2 * \ln(n)}$ arestas ou se $k_m \geq \ln(n)$, existe uma grande chance dele ser conectado.

Quase todos os grafos aleatórios com mesma quantidade n de vértices possuem aproximadamente o mesmo diâmetro D quando a probabilidade p ou a quantidade de vizinhos k_m são suficientemente grandes. Eles costumam ser dispersos, ou seja, a

vizinhança de cada vértice inclui vários novos vizinhos ainda não conhecidos entre si. O número de vértices a uma distância j de qualquer outro vértice nunca é menor que:

$$k * (k - 1)^{j-1}$$

O grau de clusterização médio de um grafo aleatório com km conexões por indivíduo é

$$\gamma_{\text{random}} = \frac{(km - 1)}{n}$$

e seu L costuma ser bastante baixo. Neste tipo de grafo, L costuma ser calculado por computador através de amostragem devido ao grande número de operações necessárias para o seu cálculo.

2.3 Grafos do tipo “connected caveman”

É um grafo extremamente organizado, altamente clusterizado, onde muitos dos vértices estão adjacentes entre si mas que ao mesmo tempo é esparso, ou seja,

$$km \ll n \text{ e } \gamma \approx 1 \text{ (Watts, 1999:102).}$$

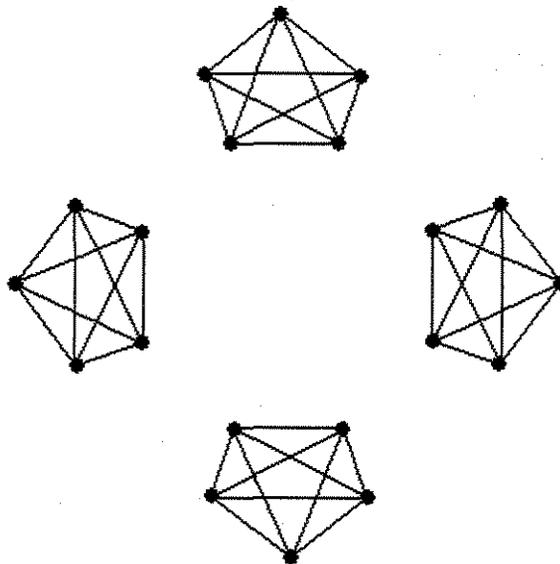


Figura 10 – Exemplo de grafo do tipo “isolated caveman”

Neste modelo, cada *cluster* possui vértices totalmente conectados entre si, com k vizinhos e portanto a quantidade de vértices por cluster é dada por :

$$n_{local} = k + 1$$

A quantidade de clusters em todo o grafo é dada por :

$$n_{global} = \frac{n}{n_{local}} = \frac{n}{(k + 1)}$$

No entanto este grafo não é conectado, conforme a Figura 10.

Para torná-lo conectado, e mantendo-se aproximadamente as mesmas propriedades de $km \ll n$ e $\gamma \approx 1$, são feitas algumas alterações em suas conexões, resultando no grafo a seguir :

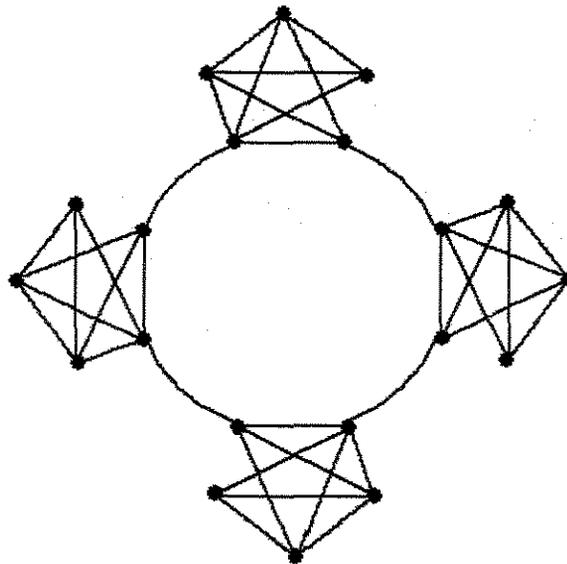


Figura 11 – Exemplo de grafo “connected caveman”

No Apêndice A é mostrado como se calcula L neste modelo de grafo.

2.4 Grafos “mundo pequeno”

As explicações a seguir foram feitas utilizando-se Watts (1999 e 2003) como referência. Grafos de “mundo pequeno” possuem as características de poucos intermediários entre quaisquer pares de nós e alta clusterização, como por exemplo na Figura 12.

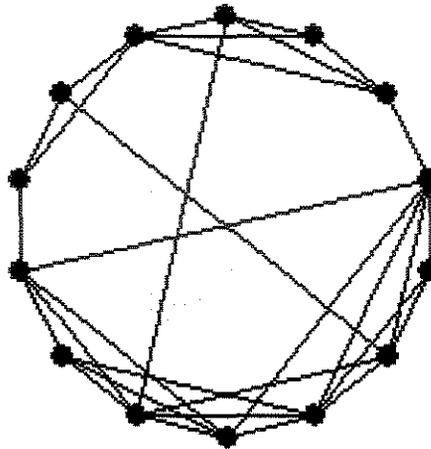


Figura 12 – Exemplo de grafo aleatório

Para tentar obtê-los, Watts fez experiências com diferentes modelos dentro de um espectro entre grafos extremamente desorganizados e extremamente organizados. Para limitar este espectro, ele utilizou em um extremo os modelos de grafos aleatórios como desorganizados, e no outro os do tipo d-lattice e connected caveman como organizados. Estas experiências foram feitas com dois tipos de categorias de grafos : grafos espaciais (com métrica) e grafos relacionais (sem métrica entre vértices).

Ele só obteve sucesso na geração de modelos de “mundos pequenos” com grafos relacionais.

Entre grafos extremamente desordenados e extremamente ordenados, Watts definiu três modelos em suas experiências :

2.4.1 Modelo α

Este modelo representa a construção de uma rede social mais realista (Watts, 1999). A forma como são geradas as conexões procura seguir o mesmo padrão observado em populações do mundo real, onde geralmente as conexões são formadas através de amigos que apresentam seus conhecidos entre si. No entanto, existe uma baixa probabilidade de novas conexões completamente aleatórias. É o modelo utilizado neste trabalho.

Watts assume que a rede pode ser representada por conexões simétricas e de mesma importância entre os elementos.

Inicialmente ele imagina um mundo da “caverna”, onde cada elemento conhece todos os membros de seu grupo, ou seja, um único grupo de conhecidos completamente conectados entre si. Ele imagina então, em um extremo um mundo com várias “cavernas” conectadas, como na Figura 11, e em outro extremo um mundo de conexões aleatórias como o da Figura 9, onde os estados das conexões existentes não influenciam nas novas conexões a serem formadas. A rede social do modelo α fica em algum lugar desconhecido entre estes mundos.

Existem infinitas possibilidades intermediárias entre estes exemplos. O mapeamento matemático destas possibilidades seria descrito pelo sistema a seguir.

2.4.1.2 Algoritmo para gerar “mundos pequenos” no modelo α :

Suponha que :

- $R_{i,j}$ é a medida da propensão de que um vértice i se conecte com um vértice j ; Seu valor será nulo se eles já forem conectados;
- $m_{i,j}$ é a quantidade de vértices adjacentes em comum entre i e j ;
- k_m é a quantidade média de vizinhos por nó no grafo;

- p é a probabilidade aleatória de existir uma conexão entre i e j quaisquer com $p \ll \langle km \rangle^{-1}$;
- α é um parâmetro configurável entre 0 e infinito que representa o grau de independência das novas conexões a serem criadas em relação as já existentes.

Temos a seguinte equação :

$$R_{i,j} = 1 \text{ se } m_{i,j} \geq km$$

$$R_{i,j} = \left[\frac{m_{i,j}}{km} \right]^\alpha * (1-p) + p \text{ se } km > m_{i,j} > 0$$

$$R_{i,j} = p \text{ se } m_{i,j} = 0$$

$$R_{i,j} = 0 \text{ se a conexão entre } i \text{ e } j \text{ já existe}$$

Gerando o grafo :

1. Gerar um grafo com n vértices não conectados;
2. Escolher um vértice i ;
3. Para cada vértice j , calcular o $R_{i,j}$;
4. Somar o $R_{i,j}$ para todos os vértices j e normalizá-los para obter

$$P_{i,j} = \frac{R_{i,j}}{\sum R_{i,j}} \text{ sendo que } \sum P_{i,j} = 1 \text{ (} P \text{ é a probabilidade do nó } i \text{ se conectar}$$

ao nó j). Dividir o intervalo $[0,1]$ em $n-1$ sub-intervalos com largura $P_{i,j}$

onde $i \neq j$, onde n é a quantidade de vértices do grafo ;

5. Sortear um número no intervalo $[0,1]$. O sub-intervalo em que ele cair pertencerá ao elemento j que se conectará a i ;
6. Conectar i ao j encontrado;
7. Repetir os passos anteriores até que a quantidade de arestas M seja :

$$\frac{(k * n)}{2}$$

8. Não escolher como novo vértice do item 1 vértices que já tenham escolhido um vizinho na rodada atual.

2.4.1.3 Variações do parâmetro α :

Se $\alpha = 0$:

Neste caso, as novas conexões serão praticamente determinadas a partir das já existentes :

Se $m_{i,j} = 0$, então $R_{i,j} = p$ onde p é um valor muito baixo;

Se $k > m_{i,j} > 0$, então $R_{i,j} = \left[\frac{m_{i,j}}{km} \right]^0 * (1 - p) + p = 1 - p + p = 1$

Se $m_{i,j} > 0$, então $R_{i,j} = 1$;

Portanto existirá uma possibilidade de que o grafo não seja conectado dependendo do valor de α . Isso pode ser evitado aumentando-se k , mas comprometeria a propriedade de dispersão ($km \ll n$) necessária no “mundo pequeno”.

Se α é próximo de zero, pouca distinção é feita pelo algoritmo de construção do grafo com relação a quantidade de vizinhos em comum entre pares de vértices. Quando α aumenta mas antes de $\left(\frac{m_{i,j}}{km} \right)^\alpha \ll p$, as arestas são criadas com maior probabilidade em pares de vértices que possuam mais de um vizinho em comum. Para alguns intervalos específicos de α , o grafo estará altamente clusterizado e com a propriedade de largura de caminho característico ainda equivalente a do grafo aleatório. Para estes intervalos, o grafo é considerado “mundo pequeno”.

Com o aumento do parâmetro α , a probabilidade de conexão entre dois vértices tenderá a um mesmo valor p , e portanto o grafo tenderá a ser aleatório, mas mesmo assim ainda existirão as seguintes diferenças :

1. O substrato será preservado, mesmo para valores altos de α .
2. O algoritmo de construção do “mundo pequeno” não sorteia as arestas de uma forma completamente independente como no grafo aleatório, ele executa algumas correlações entre as arestas.

Estas diferenças terão pouca influência para $\alpha \gg 1$ e $k \gg 1$

2.4.2 Modelo β

Representa um modelo de rede social onde se isolam as propriedades do fenômeno “mundo pequeno” encontrado no modelo α , sem características sociais como vizinhos mútuos ou clusters, ou seja, é menos realista. Este modelo utiliza uma estrutura em anel que se transforma em um grafo aleatório. Seu algoritmo começa com um grafo d-lattice, onde cada vértice possui exatamente k vizinhos ($\frac{k}{2}$ de cada lado), e então ele recria as conexões entre os vértices com uma probabilidade β da seguinte forma :

1. Cada vértice i é escolhido junto com a aresta que o conecta ao vizinho mais próximo no sentido horário $(i,i+1)$;
2. É gerado um número r aleatório. Se $r \geq \beta$, então a aresta $(i,i+1)$ é mantida. Se $r < \beta$ então a aresta $(i,i+1)$ é removida e recriada de forma a conectar i com outro vértice j de qualquer local do grafo, escolhido de forma aleatória, com exceção dos que já estão conectados.
3. Depois que cada vértice foi analisado, os passos se repetem para o vizinho seguinte mais próximo $(i,i+2)$ e assim por diante. Depois de $\frac{k}{2}$ ciclos, todas as arestas do grafo terão sido analisadas.

Para $\beta = 0$, o grafo resultante continua sendo um anel. Quando $\beta = 1$, todas as arestas são recriadas de forma aleatória, resultando em um grafo bastante próximo ao do aleatório. Para os valores intermediários, no intervalo $0 < \beta < 1$, serão obtidas diferentes características cuja interpretação é difícil, mas mesmo assim é mais clara do que a do modelo anterior.

Em ambos os modelos α e β , percebemos que para valores baixos de ambos os parâmetros existe um baixo grau de aleatoriedade, onde as novas arestas são altamente correlacionadas com as existentes ou com o substrato em anel. Em ambos os modelos as arestas tendem a ser geradas entre vértices que possuam pelo menos um vizinho em comum. Para valores altos de α e β as novas arestas tendem a ser formadas entre pares de vértices aleatórios e sem correlação com as arestas existentes.

Quando as arestas estão interligadas como uma tríade (um triângulo), elas são consideradas como arestas com $r = 2$, onde r é chamado de “range” da aresta.

O “range” significa o número de Bacon entre dois vértices na ausência da aresta entre eles. As arestas com $r > 2$ são chamadas de “shortcuts” em ambos os modelos α e β . Grafos mais clusterizados possuem uma maior quantidade de tríades.

Em um conjunto de arestas onde $M = \frac{(k * n)}{2}$, a fração de arestas do tipo “shortcut” é chamada de ϕ . Este parâmetro foi utilizado por Watts para fazer comparações entre os modelos α e β , onde ele percebeu que para pequenos valores de ϕ a largura de caminho característica era bem diferente em ambos os modelos, mas para valores maiores aproximando-se de “1”, as estatísticas dos dois modelos chegam a se tornarem equivalentes.

O modelo α é considerado mais irregular do que o modelo β , portanto os fenômenos que ocorrem no segundo são mais fáceis de serem compreendidos dependendo do tipo de análise.

2.4.3 Modelo ϕ

Este modelo foi uma tentativa de Watts em unificar as propriedades dos modelos α e β para uma melhor compreensão do fenômeno “mundo pequeno” (alta clusterização γ e baixa largura de caminho característica L).

Ao se tentar medir as propriedades de largura de caminho nos modelos anteriores em função de ϕ (fração de arestas shortcut), Watts verificou que é impossível fixar ϕ para grafos de diferentes n e k . O parâmetro ϕ só pode ser especificado como parâmetro e não explicitamente. Para resolver isto, Watts criou o modelo ϕ , que pode ser construído da seguinte forma :

1. Construir um grafo 1-lattice (anel) com n vértices onde cada um possui uma quantidade de conexões k ;
2. Especificar um valor para ϕ ;
3. Reconectar de forma aleatória as arestas do 1-lattice da mesma forma como foi proposto no modelo β , mas com uma modificação onde $\phi * \frac{(km * n)}{2}$ das arestas sejam forçadas a serem shortcuts ($r > 2$).

Para isso, seguir os seguintes passos :

- a. Selecionar um vértice u aleatoriamente;
- b. Selecionar um vizinho v aleatoriamente até que ambos tenham um vizinho em comum;
- c. Apagar a aresta (u,v) ;
- d. Escolher um vértice w aleatoriamente até que seja encontrado um que não possua vizinhos em comum com u , gerando assim uma aresta do tipo shortcut;
- e. Criar a aresta (u,w) ;

Desta forma, Watts conseguiu construir grafos do modelo ϕ com qualquer quantidade desejada de arestas do tipo “shortcut” para qualquer n e k . Isto permitiu o estudo de propriedades dos grafos com valores fixos de ϕ . Mesmo assim ele concluiu que apesar de obter propriedades interessantes, como por exemplo a complexidade de estrutura através da idéia de shortcuts, ainda não foi o suficiente para se descobrir alguma novidade no comportamento que gera o “mundo pequeno”.

2.4.4 Parâmetros que indicam chances do grafo ser “mundo pequeno”

Watts definiu o grafo “mundo pequeno” como tendo n vértices e k conexões por nó onde $L_{sw} \approx L_{random}$ mas $\gamma_{sw} \gg \gamma_{random}$ onde

$$\gamma_{random} \approx \frac{k}{n}.$$

Ele sugere comparar a sua largura de caminho característica com a do grafo aleatório e de caveman onde :

Lrandom ≈ Lsw << Lcaveman

A largura de caminho característica Lrandom é gerada para um grafo aleatório.

$$\gamma_{sw} = k_{\text{médio}} * \frac{|E(\Gamma(v))|}{(|E(\Gamma(v))| * (|E(\Gamma(v))| - 1))}$$

2.4.5 Exemplos práticos de grafos “mundo pequeno”

O ator Kevin Bacon não foi o primeiro a ser associado com o fenômeno “mundo pequeno”. Existiu antes dele um matemático chamado Paul Erdos, conhecido por ter escrito mais de 1400 papers ao longo de sua vida. Considerando-se que as pessoas que foram co-autoras destes papers receberam o número um de Erdos, as pessoas que não publicaram com ele mas que publicaram com um destes co-autores receberam o número dois de Erdos e assim por diante. A partir desta lógica, construiu-se um grafo chamado “grafo de colaboração”, cujos vértices foram os autores e as arestas representaram as suas co-autorias. Este grafo chamou a atenção dos estudiosos mas por ser incompleto acabou sendo pouco analisado. Aproveitando informações ainda incompletas disponíveis na Internet, Duncan conseguiu calcular o número de Erdos aproximado. Existem 492 autores na vizinhança de Erdos, conectados em média a 5.76 outros autores da própria vizinhança de Erdos ($\Gamma(v)$). Com estes números pode-se estimar que o índice de clusterização na vizinhança de Erdos ficou:

$$\gamma = 5.76 * \frac{492}{(492 * (492 - 1))} \approx 0.012$$

Embora pareça um número baixo, deve-se levar em consideração que Erdos possuiu mais co-autores (k) do que qualquer um da história da matemática e portanto sua vizinhança é mais ampla do que a média. Além disso, em um grafo aleatório temos:

$$\gamma_{\text{random}} \approx \frac{k}{n}$$

Se supormos um autor com o mesmo número de vizinhos e que já existiram por exemplo 100000 autores até hoje, teríamos $\gamma_{\text{random}} \approx \frac{492}{100000} = 0.00492$, que é um número bem menor do que o obtido para Erdos. O grafo de colaboração de Erdos possui

portanto alto nível de clusterização e a sua largura de caminho característica é pequena (Watts,1999), ou seja, ele pode ser um grafo do tipo “mundo pequeno”. Para se ter certeza seria necessário conhecer todas as conexões do grafo.

Os estudiosos escolheram então outros grafos onde possuíam as informações sobre todos os vértices e arestas existentes, conseguindo assim obter medições precisas. Entre eles temos os seguintes :

2.4.5.1 Grafo de Kevin Bacon

Obtido a partir da IMDb (Internet Movie Database), conforme mencionado no item sobre fundamentos de “mundo pequeno”. Este grafo foi criado considerando-se que cada autor é um vértice e que cada par de autores em um mesmo filme forma uma aresta. Foram utilizados os dados de 90% dos atores que existiram até abril de 1997, aproximadamente 225000 atores em cerca de 110000 filmes.

Obteve-se $n=225226$, $k \approx 61$, $L=3.65$ $\gamma = 0.79 + -0.02$. Utilizando-se estes parâmetros para calcular a largura de caminho característica para os grafos caveman e aleatório, obteve-se $L_{caveman} \approx 1800$ e $L_{random} \approx 3$.

A clusterização do grafo aleatório para estes parâmetros seria aproximadamente

$$\gamma_{random} \approx \frac{k}{n} \approx 0.00027, \text{ portanto este grafo é considerado um “mundo pequeno”}$$

levando-se em consideração que sua largura de caminho característica L é próxima do L_{random} e que sua clusterização γ é bem maior do que a do grafo aleatório.

2.4.5.2 Grafo da distribuição de energia a oeste dos EUA :

Ao gerar este grafo, os seguintes fatores foram considerados :

- As linhas de transmissões são consideradas como bidirecionais, ou seja, o grafo será não direcionado;
- Os nós da rede (geradores, transformadores, sub-estações,etc) são considerados como idênticos ao serem gerados os vértices do grafo;
- As linhas de transmissão são consideradas idênticas (sem pesos) para facilitar a montagem do grafo;

- Somente a rede de transmissão foi analisada, ou seja, a rede de distribuição não foi considerada.

Para este grafo, obteve-se os seguintes valores :

$$n = 4941, k = 2.67, L = 18.7 \text{ e } \gamma = 0.08$$

O valor da clusterização ficou aproximadamente 160 vezes maior do que o seu valor com parâmetros equivalentes de n e k para um grafo aleatório.

A largura de caminho característica ficou apenas 1.5 vezes superior a de um grafo aleatório com parâmetros equivalentes, conforme a tabela abaixo, portanto γ e L satisfazem as condições necessárias para um grafo “mundo pequeno” onde $L_{sw} \approx L_{rand}$ e $\gamma_{sw} \gg \gamma_{rand}$

	Grafo Distribuição de Energia	Grafo caveman	Grafo 1-lattice	Grafo aleatório
L	18.7	674	926	12.4
γ	0.08	0.65	0.3	0.0005

Tabela 1 – Comparativo da clusterização e largura de caminho característico para Distribuição de Energia no oeste dos EUA

2.4.5.3 Grafo da rede neural do sistema nervoso de uma espécie de formiga (*Caenorhabditis elegans*) :

Este grafo é um pouco menos realista porque o sistema nervoso desta formiga possui diversas irregularidades difíceis de serem reproduzidas. As principais irregularidades são as seguintes :

- Somente 302 neurônios, que é um número considerado baixo para a simulação;
- 118 classes distintas localizadas em diferentes estruturas;

- As células são interligadas por vários tipos de sinapses com sentidos diferentes e existem várias ligações entre um mesmo par de células;
- Grande variação na quantidade de ligações por célula;

Apesar destas dificuldades, Watts criou o grafo para este sistema sem se preocupar muito com os detalhes biológicos com as seguintes observações :

1. Foram ignoradas 20 células de um nervo em anel cuja conectividade ainda não foi suficientemente analisada;
2. Os 282 neurônios restantes foram considerados como idênticos;
3. As sinapses e junções entre os neurônios foram considerados como iguais;
4. A quantidade de ligações entre um mesmo par de neurônios foi simplificada para apenas considerar se eles estão ou não adjacentes;
5. As arestas foram consideradas como sendo bidirecionais.

Obteve-se $n=282$ e $k=14$, $L=2.65$, $\gamma = 0.28$:

	Grafo do sistema nervoso da C. Elegans	Grafo caveman	Grafo 1-lattice	Grafo aleatório
L	2.65	11	10.5	2.25
γ	0.28	0.98	0.69	0.050

Tabela 2 – Comparativo da clusterização e largura de caminho característico para sistema nervoso da formiga C. Elegans

A largura de caminho característica obtida é próxima a de um grafo aleatório com os mesmos parâmetros e um coeficiente de clusterização maior, indicando que o grafo possui características de “mundo pequeno”.

2.5 Resumo de L e γ

Segue abaixo um resumo das características de L e γ em cada grafo :

	Largura de caminho L	Clusterização γ
Grafos 1-lattice	com $k \geq 2$ $L = n * \frac{(n+k-2)}{(2 * k * (n-1))}$	$\gamma = 3 * \frac{(k-2)}{(4 * (k-1))}$
Grafos aleatórios	L baixo	$\gamma_{random} = \frac{(k-1)}{n}$
Grafos connected caveman	para $n \gg k \gg 1$: $L = \frac{n}{(2 * (k+1))}$	$\gamma_{caveman} = \frac{((k-2) * \gamma_a + 1 * \gamma_b + 1 * \gamma_c + 1 * \gamma_d)}{(k+1)}$ onde $\gamma_a = 1 - \frac{2}{k * (k-1)}$ $\gamma_b = 1$ $\gamma_c = 1 - \frac{2}{k}$ $\gamma_d = 1 - \frac{2}{k}$
Grafos “mundo pequeno”	$L_{random} \approx L_{sw} \ll L_{caveman}$	$\gamma_{sw} \gg \gamma_{random}$ onde $\gamma_{random} \approx \frac{k}{n}$ $\gamma_{sw} \approx \gamma_{erdos} = k_{m\u00e9dio} * \frac{ E(\Gamma(v)) }{E (\Gamma(v)) * E(\Gamma(v)) - 1}$

Tabela 3 – Resumo do cálculo da clusterização e largura de caminho característico para alguns tipos de grafos

Capítulo 3

Modelo para estudo prático

3.1 Características dos modelos de grafos utilizados

Para conseguirmos simular o comportamento de populações de “mundo aleatório” e “mundo pequeno” perante inovações, foi utilizada uma representação de tais mundos através de grafos.

As características aqui mencionadas são comuns a todos os modelos utilizados :

- **Simetria** : não foi analisado o sentido das arestas, portanto $M = \frac{(n * k)}{2}$ onde M é a quantidade de arestas do grafo G;
- **Pesos** : todas as arestas possuem o mesmo grau de influência (pesos). Esta característica não ocorre no mundo real mas foi utilizada para facilitar o cálculo e o foco no objetivo proposto neste trabalho;
- **Simplicidade** : só existe uma aresta entre cada par de vértices. Esta característica parece óbvia mas no mundo real podem existir diversos níveis de relacionamento entre dois indivíduos, como por exemplo o relacionamento de trabalho, o pessoal e o de lazer, onde cada um poderia ser uma aresta adicional ;
- **Conectividade** : todos os grafos utilizados são conexos, ou seja, não existem vértices isolados ou com apenas uma conexão.

Os grafos utilizados em cada tipo de análise possuem o mesmo valor km de quantidade média de conexões por vértice. Assim podemos avaliar a influência de outras características da estrutura social nos resultados de cada modelo.

Foi desenvolvido um modelo equivalente ao descrito no artigo de Maienhofer e Finholt (2000). Os autores analisaram o comportamento da difusão em grafos aleatórios. Tal modelo é comentado no item sobre trabalhos relacionados com o tema. O propósito é o de gerar análises de performance de difusão nos grafos de “mundo aleatório” e “mundo pequeno” e compará-las, procurando identificar as diferenças sobre o comportamento de suas populações perante uma inovação.

O modelo prático é constituído por :

- n nós onde cada um possui um número de identificação;
- km conexões médias por nó;
- *threshold* de adoção mínimo l para toda a população;
- *threshold* de adoção máximo a_{max} para toda a população;
- *threshold* de adoção aleatório a_j em cada nó j do grafo. Seu valor é sorteado no intervalo entre l e a_{max} ;
- capacidade de memorizar a quantidade de vizinhos adotantes para cada nó ao longo dos últimos NODEMEM ciclos de tempo;
- quantidade de nós conhecidos e adotantes nos últimos NODEMEM ciclos de tempo, representada pela variável d ;
- estado de adoção ou rejeição em qualquer momento durante ou após a geração do grafo representado pela variável *status*;
- quantidade inicial de inovadores, INNOVATORS;
- peso w_1 da influência da vizinhança de cada nó em sua opinião;
- peso w_2 da influência das variações de adoção ou rejeição da sua vizinhança em uma certa quantidade de ciclos.

A memória permite que seja calculada a tendência da vizinhança em adotar ou rejeitar a inovação por uma certa quantidade de ciclos. Neste modelo de simulação, depois que o potencial adotante adotar ele não poderá mais rejeitar a inovação. A quantidade inicial de inovadores é representada pelos nós que possuem menor *threshold* para adoção.

As variáveis w_1 e w_2 são utilizadas no cálculo da média de vizinhos adotantes e da sua variação nos últimos ciclos armazenados na memória, onde $d_{\text{médio}}$ representa a quantidade média de nós adotantes nos últimos NODEMEM ciclos de tempo entre os vizinhos conhecidos de um nó :

$$d_{\text{médio } j} = \frac{\sum_{i=0}^{\text{nodemem}_j - 1} d_{j,i}}{\text{nodemem}_j}$$

A variável $dd_{\text{médio}}$ representa a média da variação de $d_{\text{médio}}$ nestes últimos ciclos de tempo :

$$dd_{\text{médio } j} = \frac{\sum_{i=1}^{\text{nodemem}_j - 1} (d_{j,i} - d_{j,i-1})}{\text{nodemem}_j - 1}$$

Assim, temos que o fator de escolha e_j de cada nó será dado por :

$$e_j = w_{1,j} * d_{\text{médio},j} + w_{2,j} * dd_{\text{médio},j}$$

Se no ciclo anterior o nó não era adotante e seu fator de escolha atual for superior ao seu *threshold* de adoção a_j , então ele irá adotar a inovação, ou seja :

$$\text{Se } e_j \geq a_j \text{ então } \text{status}_j = 1$$

Os inovadores já iniciam a simulação com o status de adoção ($\text{status}_j = 1$), que permanece sempre fixo.

O grafo aleatório sorteia as arestas de forma totalmente aleatória, evitando que em um mesmo ciclo os nós que foram conectados também sorteiem uma conexão, o que

ajuda a garantir um grafo homogêneo onde todos os nós terão a mesma quantidade de vizinhos conectados.

O modelo de “mundo pequeno” utilizado foi o α , criado por Duncan Watts (cf.1999), onde o parâmetro α representa o quanto as novas conexões entre os nós serão independentes das já existentes.

O motivo da escolha deste modelo e não dos modelos β e ϕ se deve a preferência deste trabalho em seguir um modelo com mais semelhanças com a realidade de como são feitas as conexões entre os indivíduos, embora os outros modelos também sejam importantes. No modelo α as novas conexões de cada vértice são pouco aleatórias pois tendem a serem geradas a partir das já existentes, simulando a característica social existente na realidade, onde um indivíduo acaba fazendo com que seus amigos se conheçam ao apresentá-los um para o outro.

Ele utiliza a probabilidade p de conexão entre dois nós aleatórios, dada por

$$p \ll \binom{n}{2}^{-1}$$

ou seja

$$p \ll \frac{2}{(n * (n - 1))}$$

Foi introduzida no modelo uma variável aqui denominada rp (redução de p), para permitir simulações com vários níveis de redução de p , onde

$$P_{reduzido} = \frac{p}{rp}$$

Os parâmetros α , p e rp permitem a geração de grafos entre os intervalos definidos por grafos extremamente desordenados e grafos extremamente ordenados, pois sabe-se que os grafos de “mundo pequeno” estão em algum ponto deste intervalo. O modelo inclui ainda a propensão $R_{i,j}$ de conexão entre dois nós quaisquer, explicada

nos passos mencionados no modelo α descrito no item sobre grafos de “mundos pequenos”.

Para cada grafo gerado, foi medida a clusterização γ como sendo a média da quantidade de arestas na vizinhança em relação ao total de arestas possíveis na mesma para cada nó, onde a quantidade de arestas possíveis é dada por :

$$k * \frac{(k-1)}{2}$$

A fim de garantir que os grafos de “mundos aleatórios e de “mundos pequenos” seriam conexos, utilizamos grafos com substrato em anel (Watts, 1999).

Desenvolvemos um software para permitir a simulação e análise do comportamento de adoção ao longo do tempo. Utilizamos diferentes configurações de sistemas sociais com “mundo aleatório” e “mundo pequeno”. O simulador foi desenvolvido através do compilador Borland C++ Builder 3.0.

3.2 Funcionalidades do simulador

O simulador permite selecionar o tipo de arquitetura do sistema social entre o “mundo aleatório” e o “mundo pequeno” com ou sem substrato. Os grafos gerados para estas arquiteturas foram baseados no modelo proposto por Watts (1999). Outros modelos de grafos gerados pelas opções de “mundo pequeno adaptado”, com ou sem substrato, foram criados pelo autor deste trabalho. A intenção foi a de verificar como se comporta a difusão no “mundo pequeno” sem o pequeno grau de aleatoriedade utilizado por Watts em seus modelos. O simulador gera graficamente na tela o tipo de grafo selecionado, indicando com a cor “verde” os nós que se tornaram adotantes, com a cor “azul” os inovadores e com a cor “vermelha” os nós que não se tornaram adotantes. Ele também permite visualizar a curva de adotantes ao longo do tempo para o grafo desejado.

O simulador também possui opções para selecionar as características da população do sistema social, como a quantidade de indivíduos, a quantidade média de vizinhos por indivíduo, a capacidade de memória por indivíduo, a quantidade inicial de

inovadores, os *thresholds* para rejeição ou adoção da inovação pelos indivíduos e as influências dos vizinhos adotantes e das variações de adoção dos vizinhos ao longo do tempo. Também foram incluídas opções para selecionar os parâmetros de configuração do “mundo pequeno”, como o parâmetro α , a quantidade de grupos clusterizados, a probabilidade p de conexão entre dois nós quaisquer, a escolha do líder com *shortcuts* (líder global) como inovador e a escolha do líder de cada grupo clusterizado (líder local) como inovador.

Seguem nas figuras abaixo exemplos da interface do simulador :

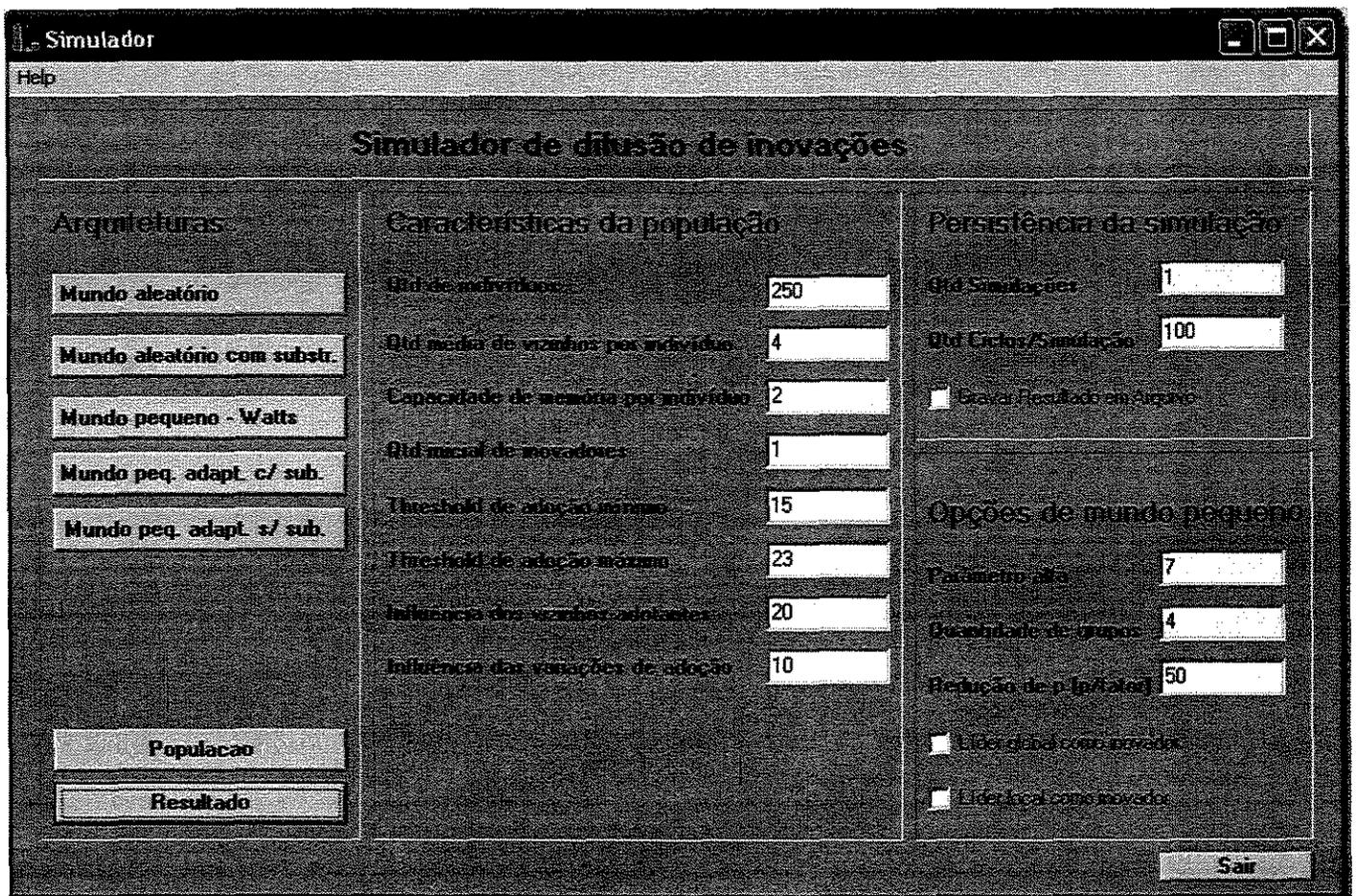


Figura 13 – Interface do software de simulação

A interface permite configurar os parâmetros da estrutura social a ser testada, assim como executar a simulação da difusão e visualizar as conexões feitas entre os indivíduos, quais são os inovadores, quais se tornaram adotantes e aqueles que rejeitaram a inovação. Além disso, permite visualizar o gráfico da curva de adoção ao longo do tempo.

Exemplos de grafos e curvas gerados :

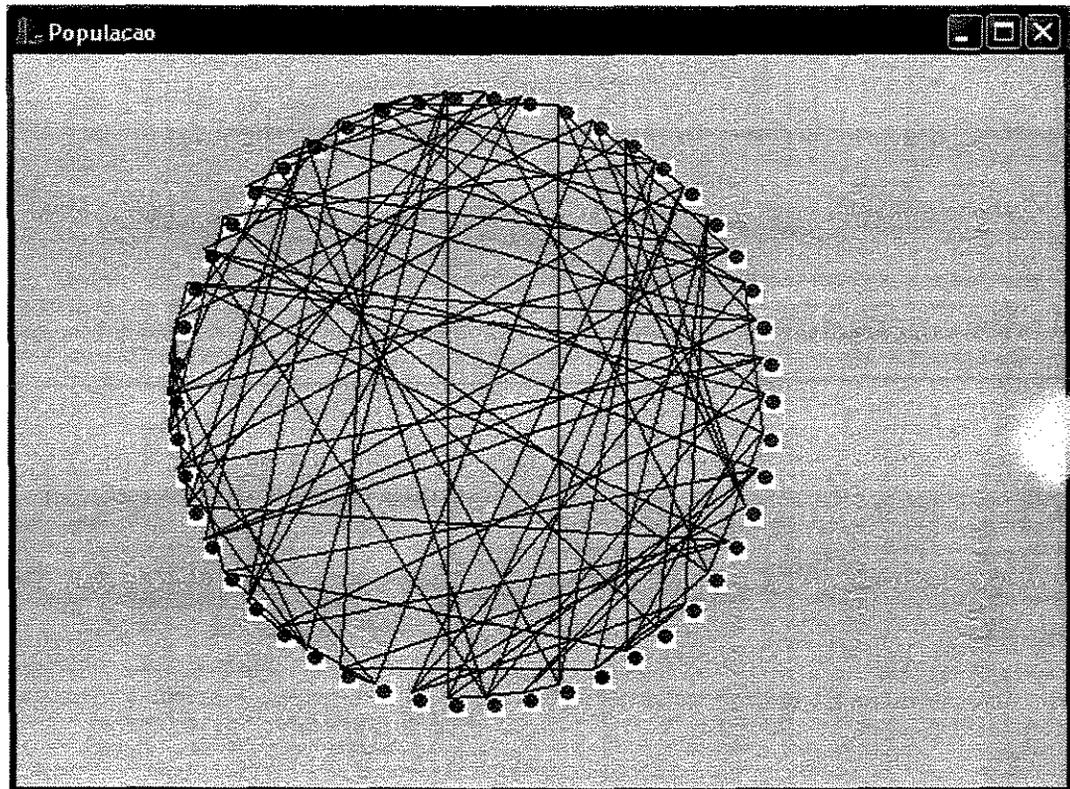


Figura 14 – Grafo aleatório obtido no simulador

Exemplo de grafo aleatório com 50 unidades, $k_{medio} = 4$ e um inovador. Este tipo de grafo possui baixa clusterização e baixa largura de caminho característica.

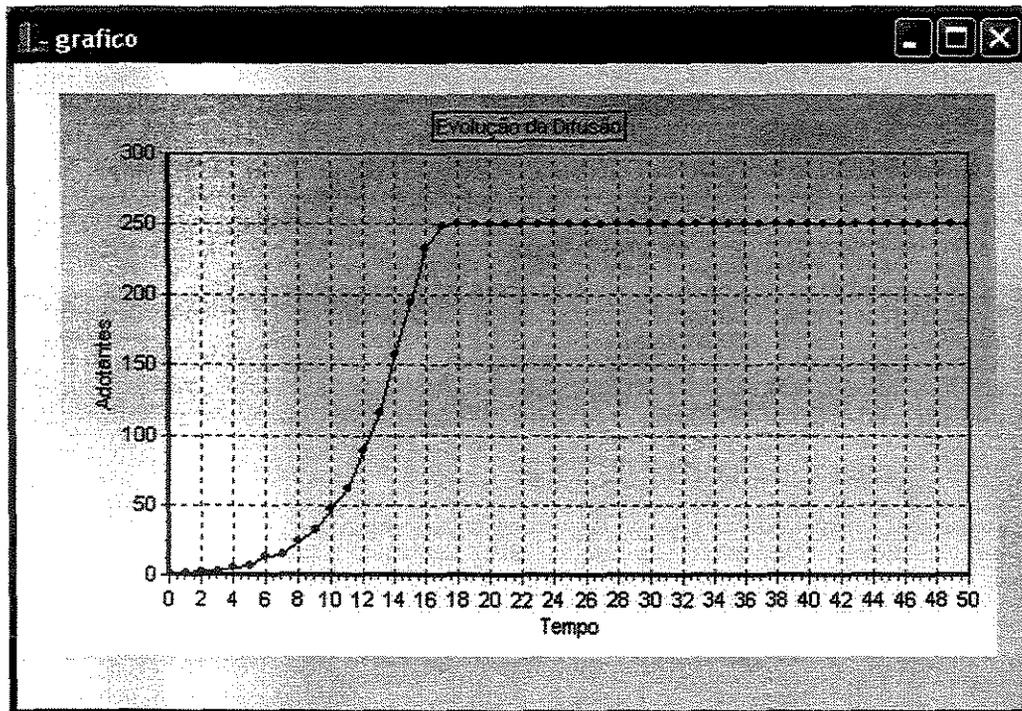


Figura 15 – Adoção ao longo do tempo - grafo aleatório

Exemplo de curva de adoção em um grafo aleatório com 250 unidades, $k_{medio} = 4$ e um inovador. Repare que a curva possui o formato em “S” encontrado em diversas análises de estudos de difusão.

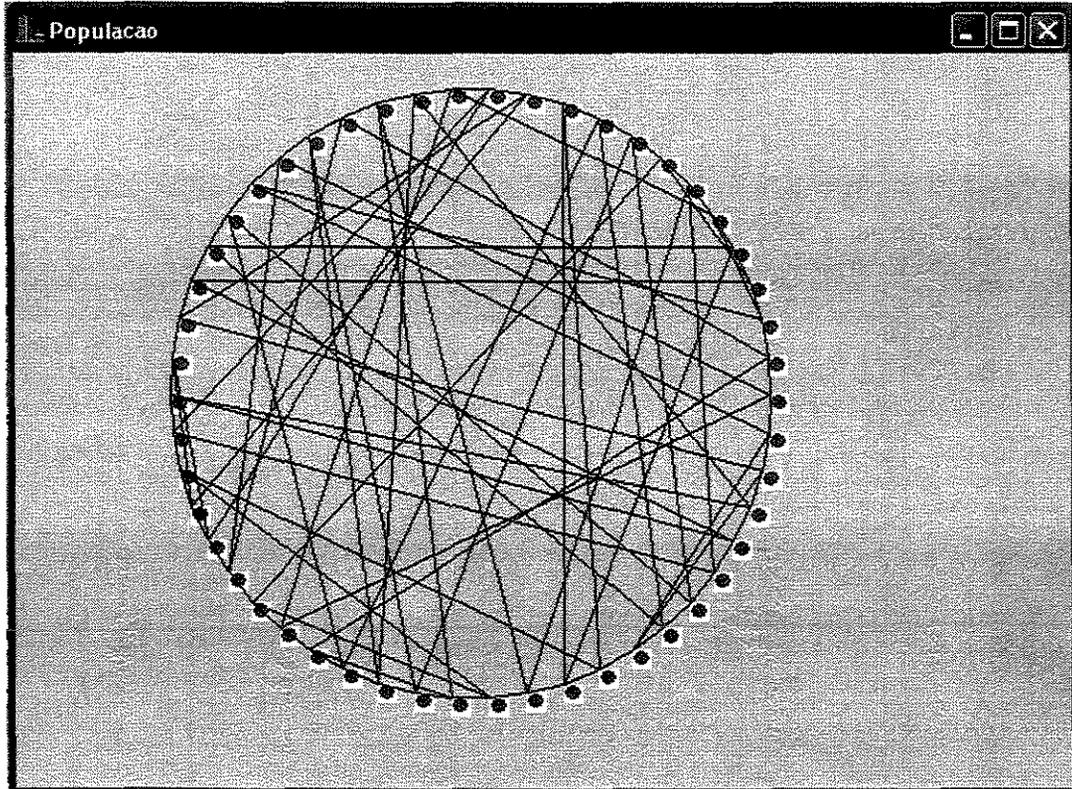


Figura 16 – Grafo aleatório com substrato obtido no simulador

Exemplo de grafo aleatório com substrato em anel, 50 unidades, $k_{medio} = 4$ e um inovador. Repare que o substrato é um anel interligando todos os nós.

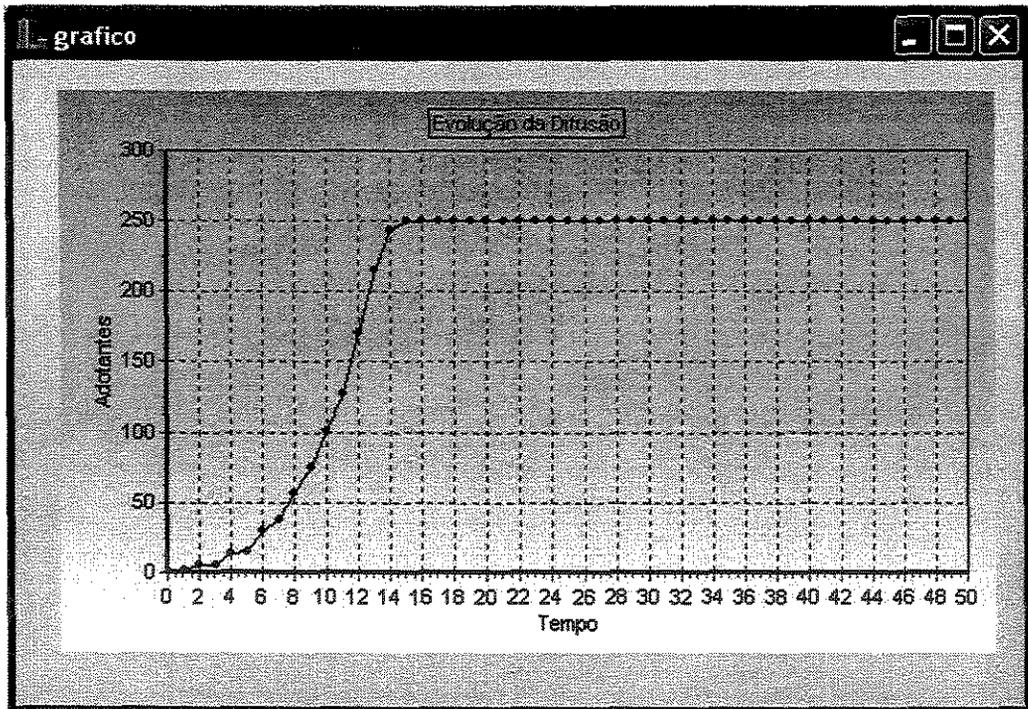


Figura 17 – Adoção ao longo do tempo - grafo aleatório com substrato

Exemplo de curva de adoção em um grafo aleatório com substrato em anel, 250 unidades, $k_{medio} = 4$ e um inovador.

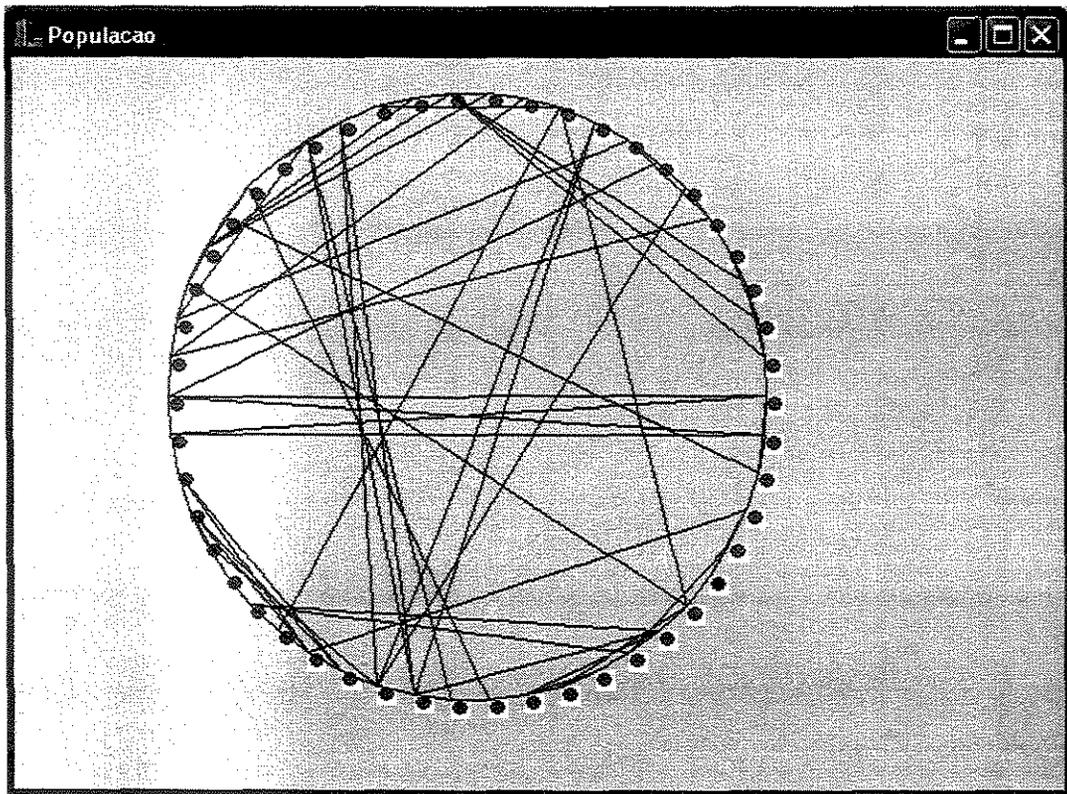


Figura 18 - Grafo “mundo pequeno” com substrato – modelo α

Exemplo de grafo “mundo pequeno” seguindo o modelo α de Watts, com substrato em anel, 50 unidades, $k_{\text{medio}} = 4$, um inovador e $\alpha=7$. Repare que existe uma maior definição entre conexões da vizinhança e conexões distantes, o que causa um aumento no índice de clusterização e uma baixa largura de caminho característica.

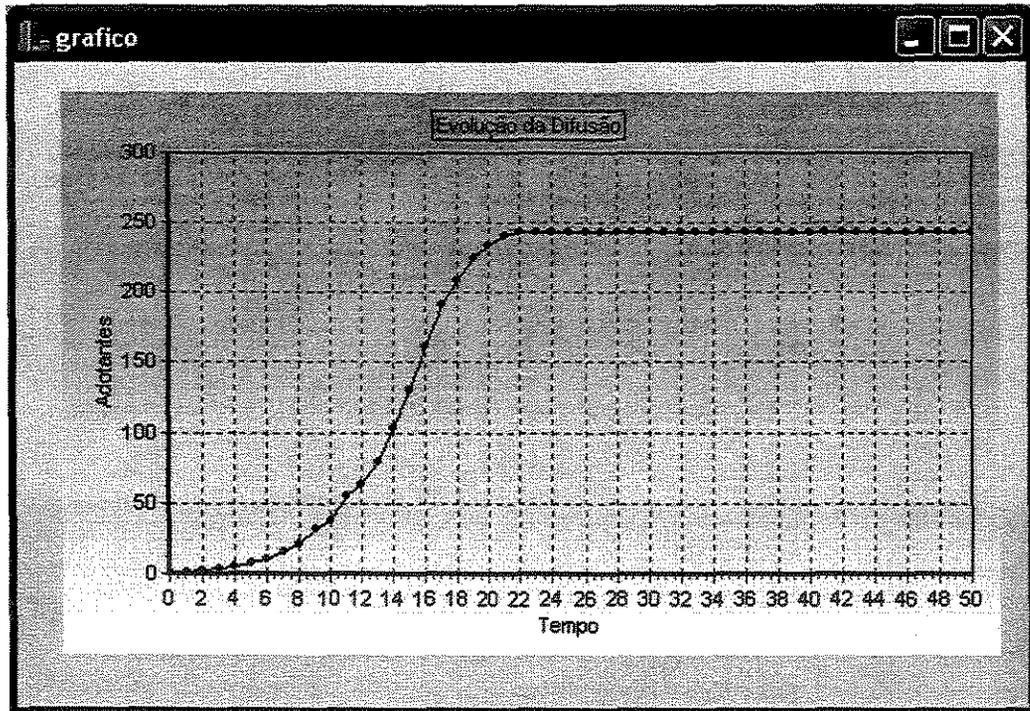


Figura 19 - Adoção ao longo do tempo - grafo “mundo pequeno” com substrato – modelo α

Exemplo de curva de adoção em um grafo “mundo pequeno” seguindo o modelo α de Watts, com substrato em anel, 250 unidades, $k_{medio} = 4$, um inovador e $\alpha=7$. Repare que neste exemplo a curva em “S” do “mundo pequeno” possui um slope maior do que a do exemplo do “mundo aleatório”.

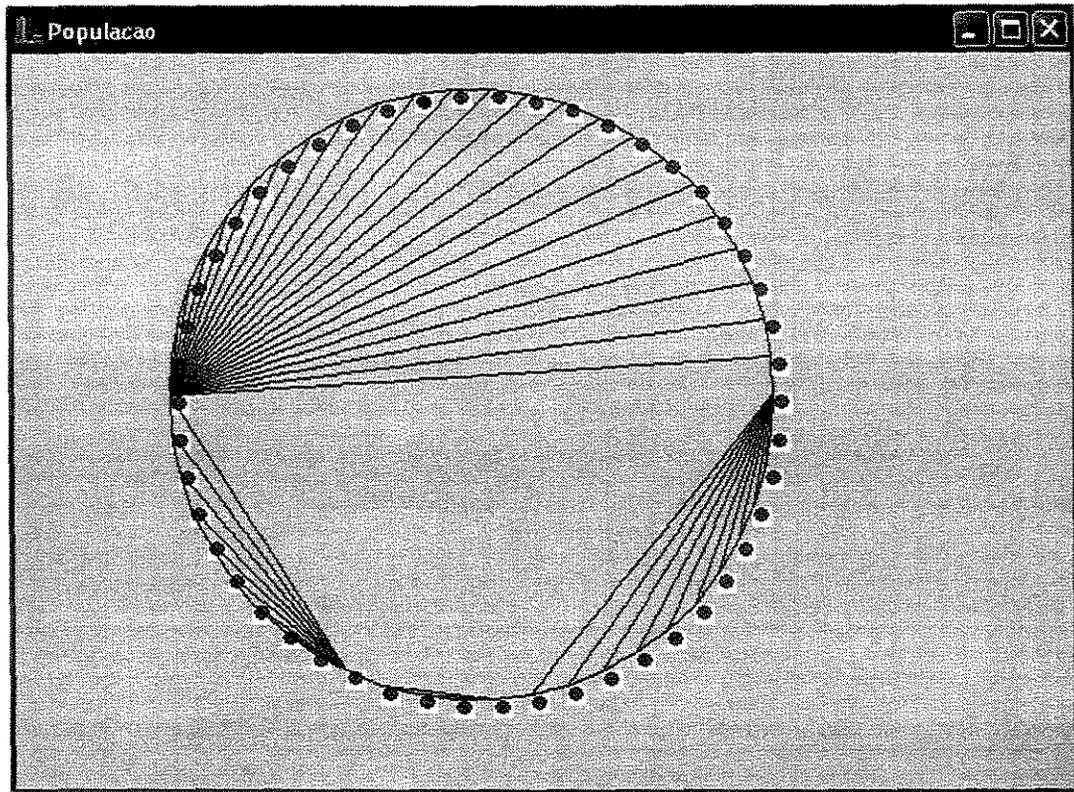


Figura 20 - Grafo “mundo pequeno” adaptado com substrato

Exemplo de grafo “mundo pequeno” com substrato em anel, 50 unidades com $k_{medio} = 4$, um inovador e 3 grupos. Este grafo foi criado pelo autor deste trabalho na tentativa de criar outros exemplos de grafos “mundo pequeno” sem aleatoriedade. Ele possui alta clusterização e baixa largura de caminho característica. Este modelo não possui links “fracos”.

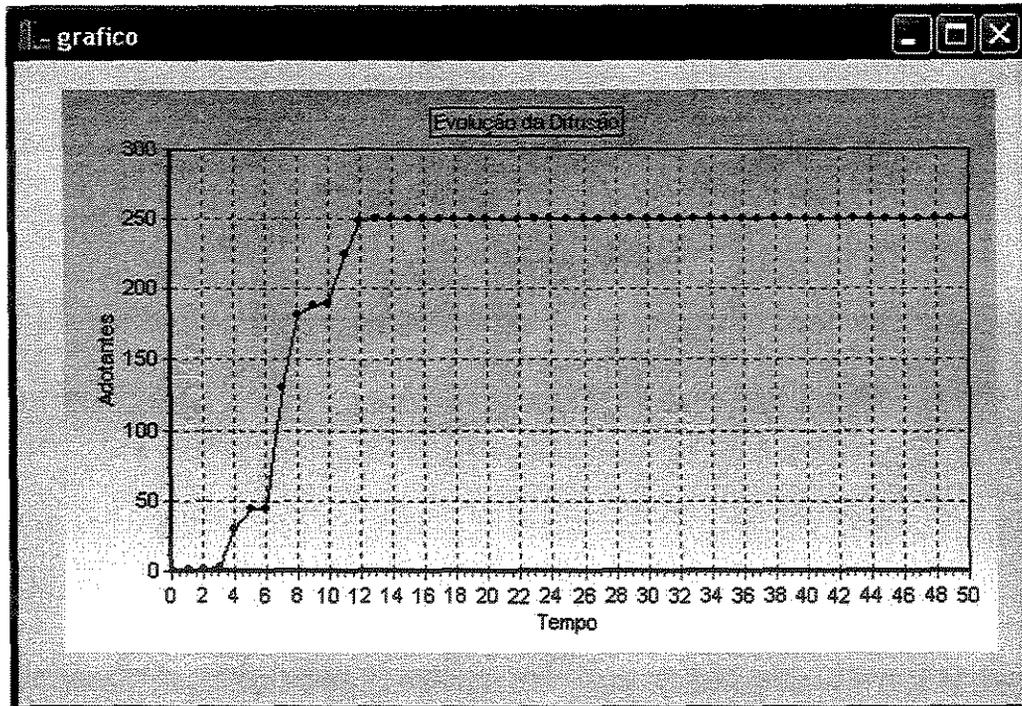


Figura 21 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado com substrato

Exemplo de curva de adoção em um grafo “mundo pequeno” adaptado com substrato em anel, 250 unidades, $k_{medio} = 4$, um inovador e três grupos clusterizados. Repare que esta curva não segue o formato padrão da curva em “S” visto nas figuras anteriores de mundo aleatório e mundo pequeno modelo α , o que é explicável pelo fato de que sua arquitetura foi feita de forma a evidenciar mais as características de clusterização e baixo L do “mundo pequeno”, e portanto não é similar a de outros modelos encontrados na natureza.

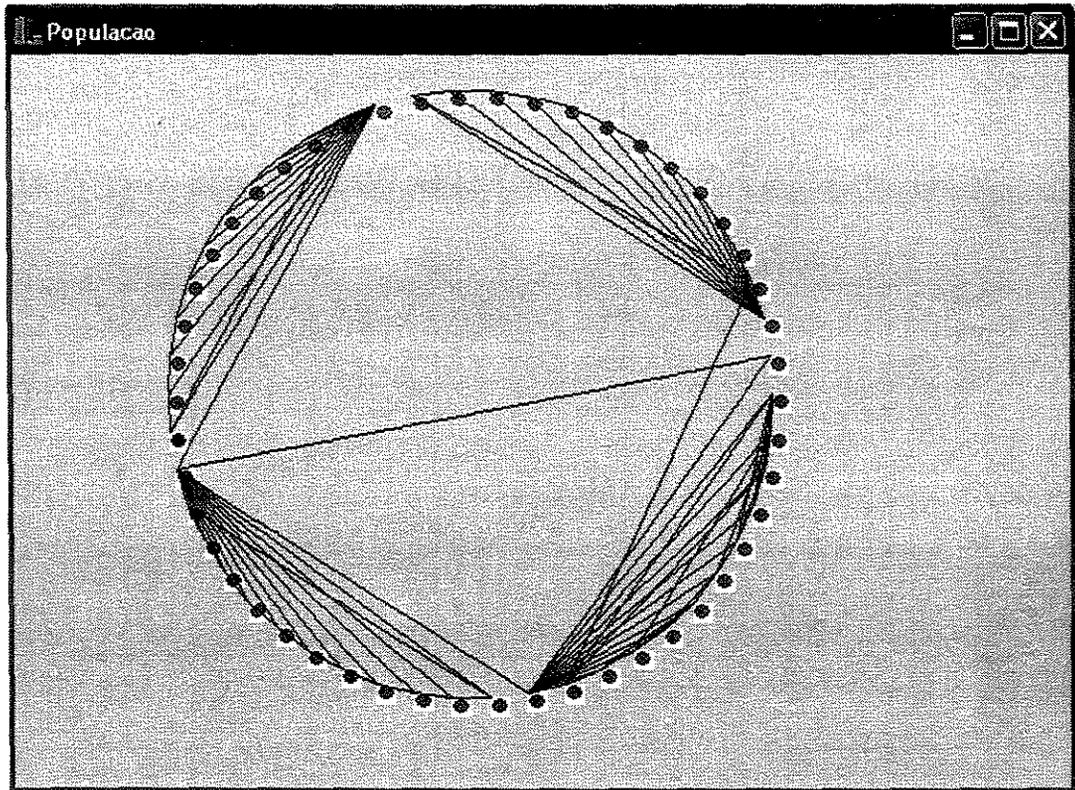


Figura 22 -Grafo “mundo pequeno” adaptado sem substrato

Exemplo de grafo “mundo pequeno” adaptado sem substrato, com 50 unidades, $k_{medio} = 4$, um inovador e quatro grupos. Este grafo também foi criado pelo autor deste trabalho na tentativa de criar outros exemplos de grafos “mundo pequeno” sem aleatoriedade. Repare que ele é parecido com o da Figura de “mundo pequeno” adaptado com substrato, mas possui links “fracos” interligando grupos distantes. Ele possui alta clusterização, baixa largura de caminho característica e não possui substrato. Este modelo foi criado conectando-se 75% dos líderes locais com os do grupo seguinte. Em seguida conectamos um membro de cada um dos últimos 25% grupos a um membro de cada um dos 25% primeiros e selecionamos o último nó do grafo como líder global. Este último nó foi conectado com 50% dos líderes locais através de links “fracos”.

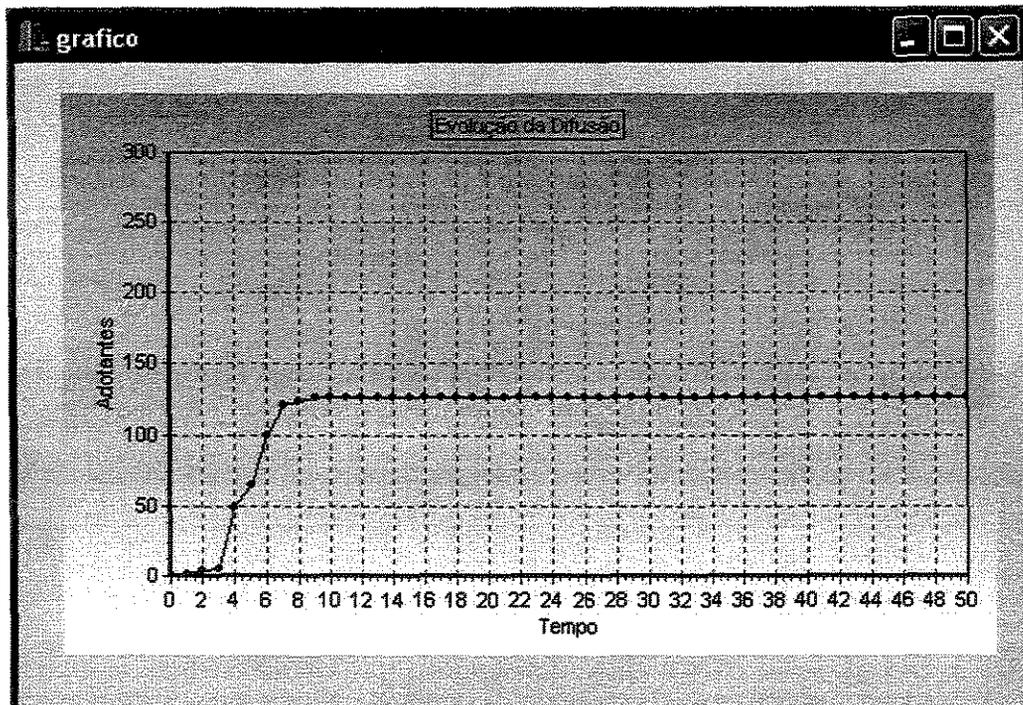


Figura 23 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado sem substrato

Exemplo de curva de adoção em um grafo “mundo pequeno” adaptado, com substrato em anel, 250 unidades, $k_{medio} = 4$, um inovador e quatro grupos. Assim como o gráfico da Figura do modelo de “mundo pequeno” adaptado com substrato, esta curva não segue o formato padrão da curva em “S” visto nas figuras de “mundo aleatório” e “mundo pequeno” modelo α , o que é explicável pelo fato de que sua arquitetura foi feita de forma a evidenciar mais as características de clusterização e baixo L do “mundo pequeno”, e portanto não é similar a de outros modelos encontrados na natureza.

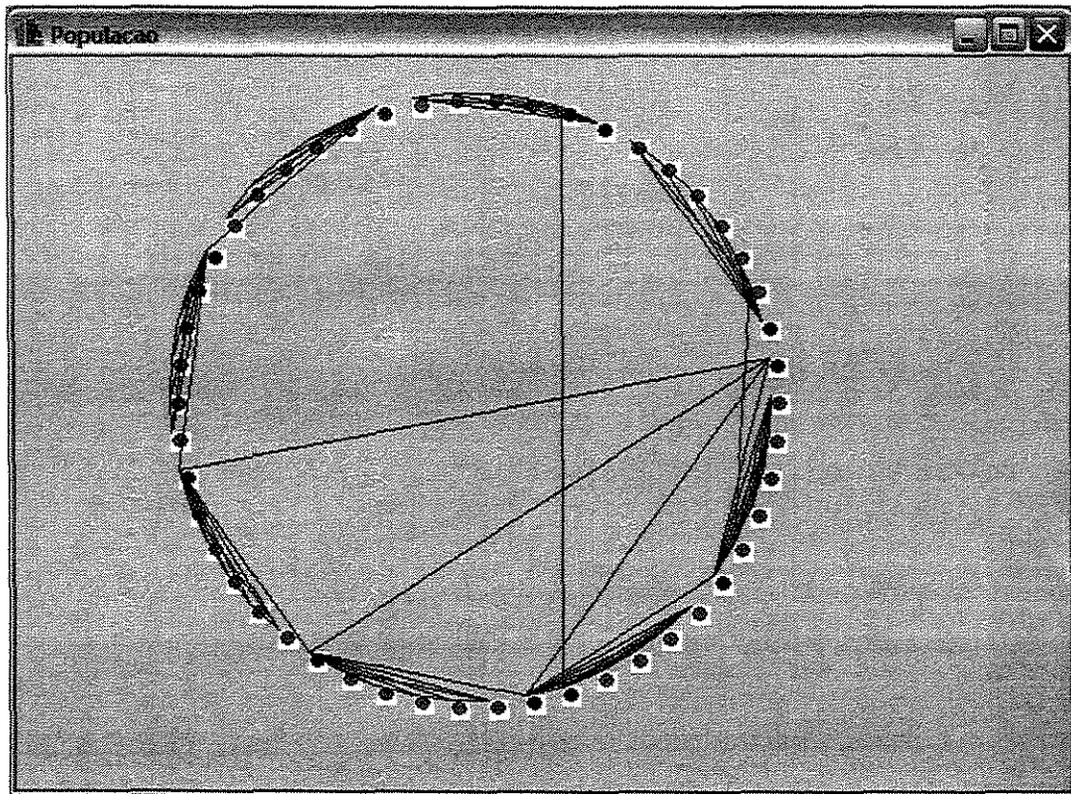


Figura 24 - Grafo “mundo pequeno” adaptado sem substrato com 8 grupos

Veja a estrutura do grafo “mundo pequeno” adaptado sem substrato com 8 grupos, 50 unidades, $k_{medio} = 4$ e líderes locais e globais como inovadores.

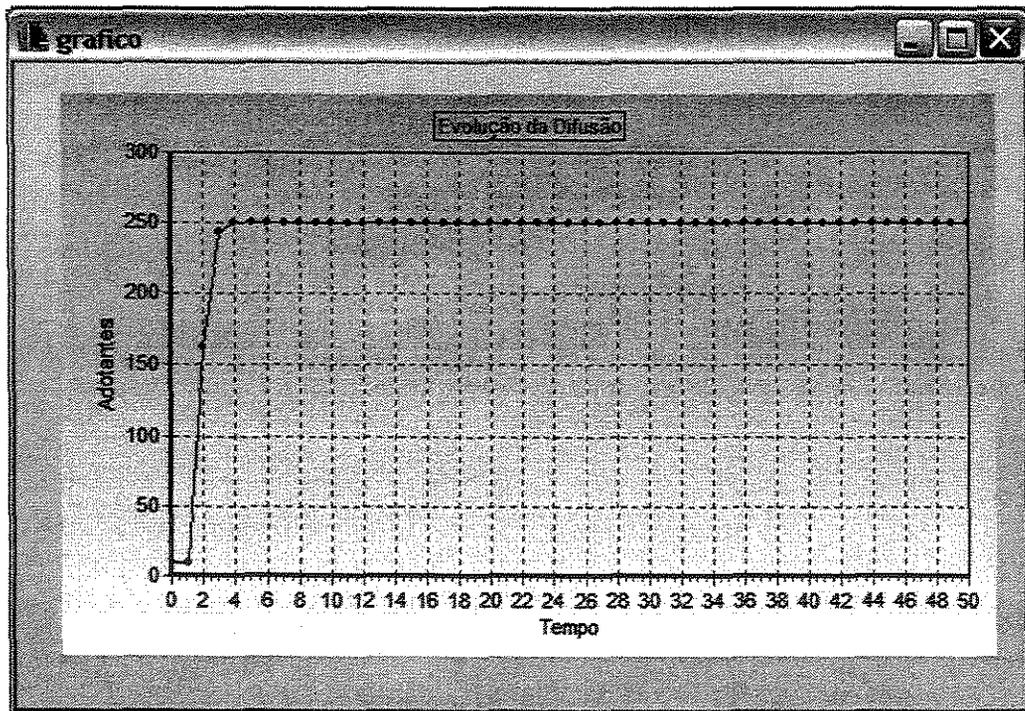


Figura 25 - Adoção ao longo do tempo - grafo “mundo pequeno” adaptado sem substrato com 8 grupos

Exemplo de curva de adoção em um grafo “mundo pequeno” adaptado sem substrato, com 250 unidades, $k_{medio} = 4$, líderes locais e globais como inovadores. Assim como o gráfico da Figura do modelo de “mundo pequeno” adaptado com substrato, esta curva não segue o formato padrão da curva em “S” visto nas figuras de “mundo aleatório” e “mundo pequeno”, o que é explicável pelo fato de que sua arquitetura foi feita de forma a evidenciar mais as características de clusterização e baixo L do “mundo pequeno”, e portanto não é similar a de outros modelos encontrados na natureza. O fato dos líderes locais e globais serem inovadores aumentou muito a performance de difusão ao longo do tempo.

Capítulo 4

Resultados do estudo prático

Para obtermos os resultados necessários, foram realizadas simulações com e sem substrato para os “mundos aleatórios” e “mundos pequenos” α e adaptado. Para cada simulação coletamos as seguintes informações :

- tipo de grafo;
- quantidade de nós;
- quantidade média de conexões por nó (k_{medio});
- número da simulação;
- coeficiente de clusterização médio;
- ciclo em que a curva ficou estável (CICLO PLAIN);
- quantidade de adotantes quando a curva ficou estável (ADOTERS PLAIN);
- quantidade de inovadores;
- coeficiente b da regressão logística da curva de adotantes ao longo do tempo até o momento da estabilização;

Nos modelos de “mundo pequeno” adaptados, simulamos a difusão com e sem a seleção de inovadores como sendo os líderes locais ou globais.

4.1 Análise de Regressão Logística

Para auxiliar na análise dos resultados foi utilizado o modelo de análise de regressão logística. Este modelo é adequado para a análise de curvas com função logística, ou seja, com formato em “S” equivalente ao mencionado no capítulo sobre os fundamentos de “difusão de inovações” na Figura 1.

Supondo que a curva de adoção de uma inovação possui uma variação influenciada pelos estados adotar (1) ou não adotar (0) de cada indivíduo, temos que a média de adotantes em toda a população é calculada pela soma de indivíduos da amostra

com estado “1” sob a quantidade total de indivíduos da população. A média de adotantes será representada pela a variável “P” e a média de não adotantes será representada pela variável Q, que vale $1 - P$. Assim, se tivermos uma amostra de 250 potenciais adotantes onde 135 adotam a tecnologia e 115 a rejeitam, teremos uma média P de $(\frac{135}{250}) = 0.54$ e $Q = 1 - 0.54 = 0.46$. Pode-se perceber que P representa a probabilidade de que um indivíduo da população sorteado aleatoriamente seja adotante. Como a proporção de adotantes ao longo da curva segue uma função logística, temos que :

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}}$$

ou

$$P = \frac{1}{1 + e^{-(a+bX)}}$$

onde a é o valor de y para $x=0$ e P é a probabilidade de um elemento possuir estado de adotante.

O valor da relação entre P e Q em um ciclo qualquer de tempo na curva é chamado de “odds”, que é calculado por :

$$odds = \frac{P}{1 - P}$$

Na análise de regressão logística, transforma-se odds em uma variação logarítmica representada por uma variável chamada “logit” cujo cálculo é feito da seguinte forma:

$$\log(odds) = \logit(P) = \ln\left(\frac{P}{1 - P}\right)$$

Desta forma, a equação para análise logística será dada por :

$$\logit(P) = a + bX$$

O logit varia linearmente com x , portanto podemos aplicar uma Regressão Linear fazendo :

$$y = a + bX$$

onde :

$$y = \text{logit}(P) = \ln\left(\frac{P}{1-P}\right)$$

e

$$P = \frac{(\text{número de adotantes})}{(\text{população de potenciais adotantes})}$$

Portanto teremos uma regressão linear com :

$$\ln\left(\frac{P}{1-P}\right) = a + bX$$

Na análise dos resultados, utilizamos a média do **coeficiente b** da regressão logística como dado comparativo entre cada tipo de simulação. Ele representa a velocidade com que a curva varia entre o seu ponto de origem e o ponto em que fica plana. Portanto, grafos com coeficientes b maiores indicam uma maior velocidade de adoção até que não apareçam novos adotantes.

4.2 Resultados e análise

Foram utilizados os símbolos definidos no item 3.1 para representarem os parâmetros utilizados nas simulações.

Percebemos que, para algumas simulações, a difusão não atingiu toda a população. Nestes casos, calculamos a percentagem de adotantes na população, descartando as amostras onde o seu valor estava inferior a 50%. Este valor de percentagem foi escolhido com base no artigo de Maienhofer e Finholt (2000), onde se considerava que somente difusões com mais de 50 % de adotantes seriam satisfatórias.

Foram feitas 50 simulações em cada modelo de estrutura social para várias configurações de parâmetros.

1. $n = 250$, $k_{\text{medio}} = 4$, NODEMEM = 2, INNOVATORS = 3, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Estes valores foram obtidos após diversas simulações, variando-se cada um deles para os modelos de grafos utilizados neste trabalho. Foram escolhidos como referência pelo fato de gerarem curvas em “S” bem definidas na maioria das simulações. Apenas lembrando, rp é o fator de redução da probabilidade p de conexão aleatória entre dois membros quaisquer, utilizado nos grafos de “mundo pequeno” modelo α .

Grafo	Clusterização média	CICLO PLAIN médio	%ADOPTERS PLAIN	b_{medio}
aleatório	0.00984	13.52	100	0.774424
aleatório com substrato	0.009	13.82	100	0.710798
“mundo pequeno” com substrato	0.354754	18.78	99.704	0.57767
“mundo pequeno” com substrato e $\alpha = 4$	0.627633	44.82	62.408	-0.01935
“mundo pequeno” com substrato e $\alpha = 10$	0.04217	13.28	100	0.859456
“mundo pequeno” com substrato e $rp = 5$	0.098537	14.32	100	0.802596
“mundo pequeno” com substrato e $rp = 500$	0.590311	41.58	88.648	0.133431
“mundo pequeno” com substrato e $rp = 5$, $\alpha = 10$	0.023183	13	100	0.809989
“mundo pequeno”	0.661969	9.66666	91.128	0.824211

adaptado com 4 grupos (pelo autor deste trabalho)		7		
---	--	---	--	--

Tabela 4 – Resultados com parâmetros de referência

O parâmetro NODEMEM=2 significa que cada nó sabe quantos vizinhos adotam no ciclo anterior e no atual.

Nos resultados deste tipo de configuração, percebemos que o modelo aleatório tende a ter uma taxa de adoção mais rápida do que a do modelo de “mundo pequeno”. Percebemos também que introduzindo um substrato em anel obtemos pouca ou nenhuma alteração nos resultados apresentados.

Em um dos resultados de $b_{\text{médio}}$ percebemos um valor negativo, indicando que a difusão foi irregular e que deveria ter sido descartada. Nesta amostra, isto ocorreu para um valor de $\alpha=4$, considerado baixo, ou seja, é um valor onde predominam novas conexões baseadas nas já existentes.

Diminuindo o parâmetro α do modelo “mundo pequeno”, percebemos que ocorre um grande aumento na sua clusterização, mas que ela não está diretamente relacionada com um aumento na velocidade de adoção. Este aumento de clusterização parece dificultar a subida da curva de adoção. Aumentando o valor de α percebemos que a clusterização diminui mas que a velocidade de adoção aumenta consideravelmente, e com nível máximo de adotantes, chegando a ser superior a do aleatório. Esta velocidade está representada no parâmetro $b_{\text{médio}}$ da regressão logística, que neste caso vale 0,8594.

Outro fato interessante é que para valores de p maiores, o grafo “mundo pequeno” modelo α , fica menos clusterizado e atinge uma velocidade de adoção próxima a do aleatório, e com nível máximo de adotantes. Se no entanto reduzirmos muito o valor de p , teremos uma considerável diminuição na velocidade de adoção, assim como maiores dificuldades na subida da curva de adoção. Aproveitando as vantagens em utilizar maior p e maior α , conseguimos obter uma velocidade de adoção superior a de todos os modelos anteriores e com nível máximo de adoção. A simulação do modelo de “mundo pequeno” adaptado que aqui desenvolvemos atingiu uma velocidade de adoção surpreendente, mas com dificuldades em obter sucesso de difusão em toda a população.

2. $n = 1000$, $k_{\text{médio}} = 4$, NODEMEM = 2, INNOVATORS = 3, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração apenas aumentamos a
quantidade de nós.

Grafo	Clusterização média	CICLO PLAIN médio	%ADOPTERS PLAIN	$b_{\text{médio}}$
aleatório	0.00206	18.24	100	0.370435
aleatório com substrato	0.00244	18.42	100	0.357534
“mundo pequeno” com substrato	0.52022704	44.42	97.50	0.193888

Tabela 5 – Resultados com aumento da quantidade de nós

Aumentando a quantidade de nós nos grafos e mantendo os mesmos parâmetros anteriores, percebemos que a velocidade de adoção do “mundo pequeno” modelo α , de Watts, diminuiu muito com relação ao de “mundo aleatório”. Isso pode ter acontecido devido à dificuldade de difusão em nós mais dispersos que não estejam em *clusters* bem definidos, que tendem a aparecer com maior frequência com o aumento da população.

3. $n = 1000$, $k_{\text{medio}} = 16$, NODEMEM = 2, INNOVATORS = 3, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração temos uma maior quantidade de nós e de conexões por nó em relação a configuração selecionada como referência.

Grafo	Clusterização média	CICLO PLAIN médio	%ADOPTERS PLAIN	b_{medio}
aleatório	0.014134	5.36	100	2.328617
aleatório com substrato	0.014033	5.24	100	2.139954
“mundo pequeno” com substrato	0.050266	5.16	100	2.10871

Tabela 6 – Resultados com aumento da quantidade de nós e de k

O aumento da quantidade de conexões por nó em populações maiores causou um surpreendente aumento na velocidade de adoção no modelo “mundo pequeno”, tornando-a ligeiramente mais rápida do que a do aleatório e com nível máximo de adoção, o que indica que este parâmetro exerce grande influência em sua performance e menos influência no modelo aleatório.

4. $n = 1000$, $k_{\text{medio}} = 4$, NODEMEM = 2, INNOVATORS = 10, $l = 15$,
 $a_{\text{max}} = 23$, $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração temos uma maior quantidade de nós e de inovadores em relação a configuração de referência.

Grafo	Clusterização média	CICLO PLAIN médio	%ADOPTERS PLAIN	b_{medio}
aleatório	0.00212	14.44	100	0.759313
aleatório com substrato	0.00219	14.58	100	0.747402
“mundo pequeno” com substrato	0.518163	35.56	97.572	0.251508

Tabela 7 – Resultados com aumento de n e de inovadores

Ao contrário do esperado, o aumento da quantidade de inovadores em populações maiores não teve o mesmo impacto causado na simulação anterior onde se aumentou k. Embora tenha melhorado a velocidade de adoção em ambos os modelos, sua influência foi pequena e manteve uma diferença de performance a favor do modelo aleatório.

5. $n = 250$, $k_{\text{medio}} = 4$, NODEMEM = 2, INNOVATORS = 3, $l = 18$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração temos um *threshold* de adoção mínimo um pouco maior do que a da configuração utilizada como referência, ou seja, a população é mais conservadora.

Grafo	Clusterização média	CICLO PLAIN médio	% simulações sem sucesso	%ADOPTERS PLAIN	b_{medio}
aleatório	0.008936	22.59574	6	100	0.283254
aleatório com substrato	0.008756	22	10	100	0.30018
“mundo pequeno” com substrato	0.352386	30.23913	8	98.5826	0.25975

Tabela 8 – Resultados com aumento do *threshold* de adoção mínimo

Aumentando o *threshold* de rejeição da população percebemos grande dificuldade na subida da curva de adoção em todos os grafos, ou seja, ocorreu uma perda de performance acima de 50% se comparada com a da primeira simulação onde $l=15$, o que indica que o grau de conservadorismo da população exerce grande influência na performance de difusão. Por isso introduzimos nesta simulação a medição da percentagem de simulações sem sucesso. No modelo “mundo pequeno” algumas difusões não obtiveram sucesso em toda a população, indicando que a sua performance é mais sensível ao grau de conservadorismo da estrutura social.

6. $n = 250$, $k_{\text{medio}} = 4$, $\text{NODEMEM} = 2$, $\text{INNOVATORS} = 3$, $l = 15$, $a_{\text{max}} = 20$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração temos um *threshold* de adoção máximo um pouco menor do que o utilizado na configuração selecionada como referência, ou seja, temos uma população menos conservadora.

Grafo	Clusterização média	CICLO PLAIN médio	% simulações sem sucesso	%ADOPTERS PLAIN	b_{medio}
aleatório	0.0086	9.64	0	100	1.349775
aleatório com substrato	0.00864	9.62	0	100	1.341447
“mundo pequeno” com substrato	0.355504	12.78	0	100	0.936187

Tabela 9 – Resultados com menor *threshold* de adoção máximo

Diminuindo o *threshold* de adoção conseguimos obter performances de difusões com melhoria expressiva no “mundo pequeno” e não tão expressiva no “mundo aleatório” em relação a configuração com $a_{\text{max}} = 23$ utilizada na primeira simulação, o que indica que este fator exerce importante influência na performance da difusão principalmente no “mundo pequeno”.

7. $n = 250$, $k_{\text{medio}} = 4$, NODEMEM = 2, INNOVATORS = 3, $l = 10$, $a_{\text{max}} = 30$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração temos um maior intervalo entre o *threshold* de adoção mínimo e o *threshold* máximo de adoção da população, ou seja, a população é mais heterogênea.

Grafo	Clusterização médio	CICLO PLAIN médio	% simulações sem sucesso	%ADOPTERS PLAIN	b_{medio}
aleatório	0.00872	16.94	0	100	0.514887
aleatório com substrato	0.00972	16.38	0	100	0.54492
“mundo pequeno” com substrato	0.351511	24.6122 4	2	99.37144	0.38609

Tabela 10 – Resultados com maior intervalo entre *thresholds* de adoção mínimo e máximo

Para uma população com maiores variações no *threshold* de adoção, percebemos uma piora na performance da difusão em ambos os modelos. Podemos destacar que no “mundo pequeno” algumas difusões não obtiveram sucesso em toda a população, o que indica que seu modelo é mais influenciável pela característica de população heterogênea.

8. $n = 250$, $k_{\text{medio}} = 4$, $\text{NODEMEM} = 2$, $\text{INNOVATORS} = 3$, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 100$, $w_2 = 0$, $rp = 50$, $\alpha = 7$. Nesta configuração aumentamos a influência dos vizinhos adotantes em relação a configuração selecionada como referência.

Grafo	Clusterização média	CICLO PLAIN médio	% simulações sem sucesso	%ADOPTERS PLAIN	b_{medio}
aleatório	0.00804	9.7	0	100	1.341208
aleatório com substrato	0.00936	9.68	0	100	1.365505
“mundo pequeno” com substrato	0.350986	13.04	0	100	0.904061

Tabela 11 – Resultados com maior influência dos vizinhos adotantes

Aumentando muito a influência dos vizinhos adotantes, conseguimos resultados muito expressivos como os da diminuição do *threshold* de adoção e, ao mesmo tempo, a performance foi bem melhor do que a da primeira simulação selecionada como referência, principalmente no “mundo pequeno”.

9. $n = 250$, $k_{\text{medio}} = 4$, **NODEMEM** = 2, **INNOVATORS** = 3, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 0$, $w_2 = 100$, $rp = 50$, $\alpha = 7$. Nesta configuração aumentamos a influência das variações de adoção em relação a configuração selecionada como referência.

Com o considerável aumento da influência da variação de adoção da vizinhança, não se obtivemos sucesso de adoção em nenhum grafo, o que indica que o estado de adoção dos vizinhos é mais importante para o sucesso da difusão do que a sua variação.

10. $n = 250$, $k_{\text{medio}} = 4$, **NODEMEM** = 10, **INNOVATORS** = 3, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta configuração aumentamos a memória por nó em relação a configuração selecionada como referência, tornando a população mais cautelosa em adotar uma inovação.

Grafo	Clusterização média	CICLO PLAIN médio
aleatório	0.00992	63.1
aleatório com substrato	0.00868	65.4
“mundo pequeno” com substrato	0.354319	90.9

Tabela 12 – Resultados com aumento da memória de cada nó

Com o aumento da memória de cada nó, percebemos que a velocidade de adoção caiu de forma bastante expressiva em ambos os modelos e que a diferença de performance entre ambos ficou ainda mais evidente.

11. $n = 250$, $k_{\text{médio}} = 4$, NODEMEM = 2, INNOVATORS = 5, $l = 15$, $a_{\text{max}} = 23$,
 $w_1 = 20$, $w_2 = 10$, $rp = 50$, $\alpha = 7$. Nesta última configuração, aumentamos a quantidade de inovadores para cinco e comparamos os modelos de “mundo aleatório” e “mundo pequeno” com os de “mundo pequeno” adaptados e com as opções de escolha de líderes como inovadores.

Grafo	Clusterização média	CICLO PLAIN médio	% simulações sem sucesso	%ADOPT ERS PLAIN	$b_{\text{médio}}$
aleatório	0.00776	11.98	0	100	0.923949
aleatório com substrato	0.0092	12.06	0	100	0.938031
“mundo pequeno” com substrato	0.351059	16.7	0	99.888	0.605422
“mundo pequeno” adaptado com 4 grupos e 5 inovadores aleatórios (pelo autor deste trabalho)	0.661969	8.72	0	95.992	1.25342
“mundo pequeno” adaptado com 4 grupos e o líder de cada um como inovador. Foi incluído um líder aleatório. (pelo autor deste trabalho)	0.661969	4.86	0	100	3.173282
“mundo pequeno” adaptado com 4 grupos, líder global como inovador e outros 4 inovadores aleatórios. (pelo autor deste trabalho)	0.661969	8.16	0	93.016	1.089766
“mundo pequeno” adaptado com 4 grupos e	0.661969	4.68	0	100	3.295862

o líder de cada um como inovador, líder global também como inovador (pelo autor deste trabalho)					
---	--	--	--	--	--

Tabela 13 – Resultados comparando-se com os modelos de “mundo pequeno” criados manualmente

Percebemos que existe sempre uma tendência da velocidade de adoção do “mundo aleatório” ser maior do que a do “mundo pequeno” modelo α , mas ela é expressivamente inferior à do “mundo pequeno” adaptado, onde foi retirada a aleatoriedade e enfatizadas as propriedades de clusterização e de largura de caminho característico. Selecionando o líder de cada cluster como inovador, alcançamos resultados surpreendentes na performance de adoção do modelo “mundo pequeno” adaptado. Ao contrário do esperado, selecionando também o líder global (ligações “fracas”) como inovador, neste modelo obtivemos pouca diferença nos resultados, indicando que os líderes locais exercem maior influência na performance da difusão se selecionados como inovadores.

Portanto, podemos concluir que, só serão alcançadas melhores performances de velocidade de adoção e de sucesso de adoção em modelos específicos de “mundo pequeno”, que são difíceis de serem elaborados. O modelo de “mundo aleatório” possui uma boa performance e é fácil de ser desenvolvido, mas, dependendo de onde for aplicado, a sua escolha pode comprometer alguma propriedade necessária à estrutura social.

4.3 Análise de Teste-t

O Teste-t permite comparar duas amostras para se verificar se as suas características são equivalentes ou diferentes com relação a uma propriedade. Utilizamos a função de Teste-t do módulo de cálculo do Excel.

Foram selecionados os modelos de “mundo aleatório” e “mundo pequeno” α de Watts com melhor performance para uma configuração de $n = 250$, $k = 4$, $NODEMEM = 2$, $INNOVATORS = 3$, $l = 15$, $a = 23$, $w_1 = 20$, $w_2 = 10$, $rp = 5$, $\alpha = 10$, encontrados no item 1 dos resultados.

Nesta análise comparamos os resultados obtidos para o CICLO PLAIN em ambos os modelos. Para a hipótese de que as amostras de CICLO PLAIN das 50 simulações de cada modelo sejam equivalentes assumindo que possuem variâncias diferentes, com uma confiança de 5%, obtivemos $P(T \leq t)$ de 0.026473 inferior a 0.05, ou seja, o Teste-t detectou diferenças entre estes parâmetros nas 2 amostras, conforme a tabela abaixo :

	<i>Grafo Aleatório</i>	<i>“mundo pequeno” com $rp=5$, $\alpha=10$</i>
Mean	13.52	13
Variance	2.050612245	1.469387755
Observations	50	50
Hypothesized Mean Difference	0	
Df	95	
t Stat	1.95982374	
P(T<=t) one-tail	0.026473421	
t Critical one-tail	1.661051101	
P(T<=t) two-tail	0.052946841	
t Critical two-tail	1.985249583	

Tabela 14 – Resultados do Teste-t

Estes resultados indicam que o ponto onde a curva se torna plana tende a ser diferente entre os modelos. Isto indica que em algum momento durante a difusão ocorre alguma variação de velocidade entre os modelos que causa uma variação razoável do ciclo em que a curva se torna plana. Percebemos também que a variância do CICLO PLAIN no “mundo aleatório” é sensivelmente superior a do “mundo pequeno”, o que

indica um maior nível de ruído em sua curva e uma maior estabilidade e previsibilidade no segundo modelo.

Capítulo 5

Conclusões

Percebemos que a performance da “difusão de inovações” no grafo “mundo pequeno” não costuma ser superior à do grafo “mundo aleatório”, com exceção de alguns modelos específicos encontrados, o que contraria o que esperávamos intuitivamente. Podem existir mais propriedades a serem consideradas na definição do “mundo pequeno”, além das conhecidas como clusterização e largura de caminho característico. Um exemplo poderia ser a curva de distribuição da população com relação à quantidade de conexões por nó. É interessante a percepção de que, no grafo aleatório, ocorreu maior ruído nos resultados obtidos para o ciclo em que a curva de adoção se estabiliza em relação aos encontrados no grafo de “mundo pequeno”. Isto indica que o “mundo aleatório” é um pouco menos previsível.

Também foi importante a observação de que pequenas variações na influência de algumas características da população causaram grandes variações positivas ou negativas nos resultados, principalmente no “mundo pequeno”, indicando sua maior sensibilidade. O modelo adaptado de “mundo pequeno”, onde retiramos a aleatoriedade das conexões e isolamos as características de largura de caminho característica e clusterização, mostrou-se o mais sensível, porém a sua performance foi a melhor.

A seguir apresentaremos as características que mais influenciaram no comportamento da difusão.

Melhoria da performance :

- aumento da influência da vizinhança de cada nó;
- quantidade média de conexões por nó;
- clusterização;
- aumento do grau de independência das novas conexões em relação às anteriores (parâmetro α);

- aumento da probabilidade aleatória de conexão entre dois membros quaisquer (parâmetro p);
- seleção de membros inovadores como líderes locais;
- eliminação da aleatoriedade e maior evidência das propriedades de largura de caminho característica e clusterização;

Piora da performance :

- aumento da influência da variação de adoção da vizinhança de cada nó;
- aumento de memória;
- aumento da heterogeneidade da população;
- aumento do conservadorismo da população;
- aumento exagerado da clusterização

O aumento da clusterização, por exemplo, não representa uma melhor performance de adoção e, em certos casos, é até capaz de prejudicar a performance. A quantidade média de conexões pode melhorar muito a performance do “mundo pequeno”, mas possui menos benefícios no modelo aleatório. Populações mais heterogêneas tornam mais difícil a difusão, principalmente no modelo de “mundo pequeno”. Outra característica interessante que percebemos é o fato de que variações de adotantes na vizinhança causam menos influência do que o estado de adoção em ambos os modelos.

Assim, podemos concluir que, embora a performance do modelo aleatório seja geralmente superior, a escolha criteriosa do modelo “mundo pequeno” pode ser vantajosa. Isto se deve ao fato de possuir características em sua estrutura social que são bastante frequentes na natureza (efeito mais “natural”) e que não são encontradas no modelo aleatório. Além disso, possui performance de difusão ainda considerada boa em relação ao aleatório. Mas deve-se atentar para a sua maior susceptibilidade às características da estrutura social.

A seleção de líderes pode tornar a performance do modelo “mundo pequeno” ainda mais eficiente do que o modelo aleatório.

Esperamos que o que foi feito neste trabalho, possa de alguma forma, auxiliar nos estudos para aplicações de “mundo pequeno” em diversas áreas e no aperfeiçoamento dos estudos de “difusão de inovações”.

5.1 Desdobramentos futuros desta pesquisa

As análises indicaram que o ciclo em que as curvas de adoção se tornam estáveis, no modelo de “mundo aleatório” e de “mundo pequeno”, possui características diferentes. Isto sugere que seria interessante um estudo mais detalhado do comportamento da curva de adoção. É possível que em certos momentos a curva possua maior velocidade em um modelo do que no outro, que é um assunto que poderia ser aprofundado em trabalhos futuros.

Em trabalhos futuros poderiam ser analisados também os modelos β e ϕ do “mundo pequeno” para verificar se as características detectadas neste trabalho também se aplicam e se a sua performance também tende a ser inferior à do “mundo aleatório”.

Igualmente de interesse seria a análise do “*power law*” em todos os modelos de “mundo pequeno” de Watts e no modelo aleatório (Mihajlo, Annexstein e Berman, 2001; Amaral, Scala, Barthélemy e Stanley, 2000).

Apêndice A

Cálculo de L em um grafo “Connected Caveman”

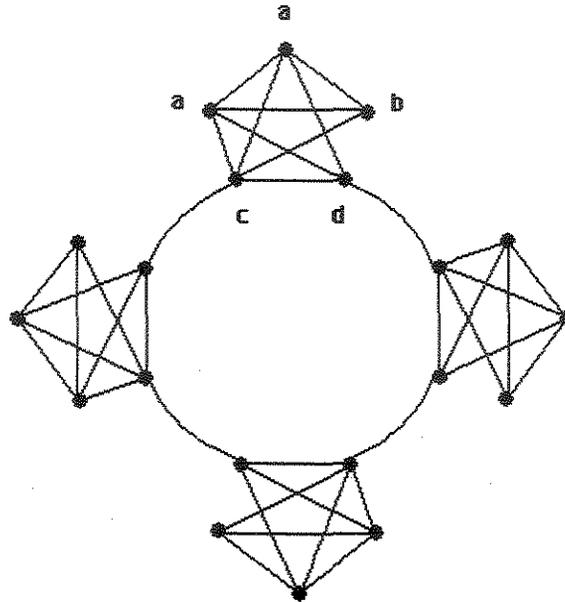


Figura 26 – Tipos de vértices do Grafo aproximado do “connected caveman”

Neste modelo aproximado, existem variações na quantidade de conexões de cada unidade pertencente aos clusters, variações estas que analisaremos para entender como Watts calculou L :

Vértices do tipo a : k-2 vértices por cluster estão conectados a todos os outros do próprio cluster e 1 aresta não existe na vizinhança :

Como a clusterização γ é dada pela quantidade de arestas existentes sobre a quantidade de arestas que poderiam existir, temos que :

Combinação de k arestas 2 a 2 vale

$$\langle k \rangle_2 = \frac{k!}{2! * (k-2)!} = \frac{k * (k-1) * (k-2)!}{2 * (k-2)!} = \frac{k * (k-1)}{2}$$

portanto

$$\gamma_a = \frac{\langle k \rangle - 1}{\langle k \rangle} = \frac{\frac{k * (k-1)}{2} - 1}{\frac{k * (k-1)}{2}} = 1 - \frac{2}{k * (k-1)}$$

Vértices tipo b : 1 vértice em cada cluster possui k-1 vizinhos. Todos os seus vizinhos estão conectados entre si portanto $\gamma_b = 1$

Vértices tipo c : 1 vértice em cada cluster possui k+1 vizinhos. Ele se conecta ao vértice tipo d de outros clusters e portanto perde k arestas de sua outra vizinhança além da que já perde em seu próprio cluster.

Portanto

$$\gamma_c = \frac{\langle k+1 \rangle - k - 1}{\langle k+1 \rangle} = \frac{\frac{(k+1) * k}{2} - k - 1}{\frac{(k+1) * k}{2}} = 1 - \frac{2}{k}$$

Vértices tipo d : 1 vértice em cada cluster está conectado ao vértice do tipo c de outro cluster e portanto perde k-1 vizinhos de outro cluster.

$$\text{Portanto } \gamma_d = \frac{\langle k \rangle - (k-1)}{\langle k \rangle} = \frac{\frac{k * (k-1)}{2} - (k-1)}{\frac{k * (k-1)}{2}} = 1 - \frac{2}{k}$$

Portanto, a clusterização ponderada total do grafo "Connected Caveman" será :

$$\begin{aligned} \gamma_{cc} &= (\text{quantidade}_a * \gamma_a + \text{quantidade}_b * \gamma_b + \text{quantidade}_c * \gamma_c + \text{quantidade}_d * \gamma_d) / \\ &\quad (\text{quantidade}_a + \text{quantidade}_b + \text{quantidade}_c + \text{quantidade}_d) \\ &= \frac{(k-2) * \gamma_a + 1 * \gamma_b + 1 * \gamma_c + 1 * \gamma_d}{(k-2) + 1 + 1 + 1} \\ &= \frac{(k-2) * \gamma_a + 1 * \gamma_b + 1 * \gamma_c + 1 * \gamma_d}{(k+1)} \end{aligned}$$

A Largura de caminho característica L deste grafo será calculada em função da distância média entre 2 vértices do mesmo cluster (d_{local}) com L_{local} e a distância média entre vértices em diferentes clusters (d_{global}) com L_{global} .

Como em cada cluster existem 2 unidades (b e d) cuja distância d_{ij} pode atingir o valor 2 e as outras possuem $d_{ij}=1$, temos :

$$\begin{aligned}
 d_{local} = L_{local} &= \left(\frac{\binom{k+1}{2} - 2 \cdot 1 + 2 \cdot 2}{\binom{k+1}{2}} \right) \\
 &= \frac{\frac{(k+1) \cdot k}{2} - 2 + 4}{\frac{(k+1) \cdot k}{2}} \\
 &= 1 + \frac{4}{(k+1) \cdot k}
 \end{aligned}$$

Com isso percebemos que para $k \gg 1$ teremos $d_{local} \approx 1$

A distância média d_{global} está em função de L_{local} e L_{global} .

Do tópico sobre Grafos lattices sabe-se que L_{global} para um anel é dado por $L_{anel} = \frac{n \cdot (n + k - 2)}{2 \cdot k \cdot (n - 1)}$, A quantidade de clusters no grafo aproximado do “connected

caveman” é dada por $n_{global} = \frac{n}{k+1}$, portanto substituindo-se n por $\frac{n}{k+1}$ teremos :

$$L_{global} = \frac{\frac{n}{(k+1)^2}}{4 \cdot \frac{n}{k+1} - 1}$$

Para se calcular d_{global} , precisa-se entender que o caminho entre 2 vértices de diferentes clusters é dado por :

1. Número de arestas para sair do cluster (L_{local});
2. Número de arestas entre os clusters, que no caso é uma global e 1 local para $L_{global} - 1$ clusters.
3. Número de arestas para chegar até o cluster $L_{local} + 1$.

Ou seja, $d_{global} = L_{local} + 2(L_{global} - 1) + (1 + L_{local})$

que para $n \gg k \gg 1$ resulta em $\frac{n}{2 * (k + 1)}$

O cálculo da largura de caminho característica “connected caveman” (L_{cc}) será em função da quantidade de arestas que pertencem ao mesmo cluster (N_{local}) e da quantidade que não pertence ao mesmo cluster (N_{global}).

N_{local} = combinação de arestas por cluster * quantidade de clusters

$$= \frac{(k + 1) * k}{2} * \frac{n}{k + 1} = \frac{n * k}{2}$$

$$N_{total} = \frac{n * (n - 1)}{2}$$

$$N_{global} = N_{total} - N_{local} = \frac{n * (n - 1)}{2} - n * \frac{k}{2} = \frac{n * (n - 1 - k)}{2}$$

Portanto L_{cc} será a distância média entre todos os pares de vértices :

$$L_{cc} = \frac{N_{local} * d_{local} + N_{global} * d_{global}}{N_{total}}$$

que para $n \gg k \gg 1$ resulta em $\frac{n}{2 * (k + 1)}$

Este parâmetro é utilizado para comparar outros grafos com o aproximado de “connected caveman”.

Bibliografia

AMARAL, L.A.N., SCALA, A., BARTHÉLÉMY, M. e STANLEY, H.E., *Classes of Small-World networks*, Center for Polymer Studies and Department of Physics, Boston University, Boston, July 2000 –
<http://polymer.bu.edu/~amaral/Papers/pnas00a.pdf>

BUCHANAN, Mark, *Nexus – Small Worlds and the Groundbreaking Science of Networks*, New York, Norton & Company, 2002

BUSSAB, Wilton de Oliveira, *Análise de Variância e de Regressão*, São Paulo, Atual Editora, maio de 1986

GOLDBERG, Debra S. e ROTH, Frederick P., *Assessing experimentally derived interactions*, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, September 2002 -
<http://www.pnas.org/cgi/reprint/100/8/4372.pdf>

JOBBER, David, *Principles and Practice of Marketing*, Second Edition, London, McGraw-Hill Publishing Company, May 1998

KLEINBERG, Jon, *The Small-World Phenomenon: An Algorithmic Perspective*, Department of Computer Science, Cornell University, Ithaca, October 1999 -
<http://www.ti.inf.ethz.ch/pw/teaching/webalgs/kleinberg.pdf>

LAGO-FERNÁNDEZ, Luis F., HUERTA, Ramón, CORBACHO, Fernando, SIGÜENZAL Juan A., *Fast Response and Temporal Coherent Oscillations in Small-World Networks*, E.T.S de Ingeniería Informática, Universidad Autónoma de Madrid, Madrid, Spain, July 1999 –
<http://inls.ucsd.edu/~huerta/papers/luis.pdf>

MAIENHOFER, Dirk and FINHOLT, Thomas, *Finding Optimal Targets for Change Agents: A computer Simulation of Innovation Diffusion*, University of Michigan. 2002
<http://www.springerlink.com/app/home/contribution.asp?wasp=a8ffb0d11d3f4f36af9d570a282a8a54&referrer=parent&backto=issue,1,5;journal,8,31;linkingpublicationresults,1:102865,1>

MIHAJLO, A., ANNEXSTEIN, Fred S. e BERMAN, Kenneth A., *Modeling Peer-To-Peer Network Topologies Through Small-World Models and Power Laws*, ECECS Department, University of Cincinnati, Cincinnati, November 2001
<http://www.telfor.org.yu/telfor2001/radovi/2-14.pdf>

PORTER, Michael E., *Competitive Strategy*, New York, Free Press, 1980

RICHERS, Raimar, *Surfando as Ondas do Mercado*, 5ª. Ed, Editora RR & CA, São Paulo, agosto de 1997

ROGERS, Everett M., *Diffusion of Innovations*, 3rd Edition, New York, Free Press, 1983

ROGERS, Everett M., *Diffusion of Innovations*, 5th Edition, New York, Free Press, 2003.

STROGATZ, Steven H., *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering*, Massachusetts, Westview Press, 1994

WATTS, Duncan J., *Small Worlds – The Dynamics of Networks between Order and Randomness*, Princeton, New Jersey, Princeton University Press, 1999

WATTS, Duncan J., *Six Degrees – The Science of a Connected Age*, New York, Norton & Company, 2003

WATTS, Duncan J., STROGATZ, Steven H., Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York, 1998